



TITLE:

最大エントロピー法を用いた水文量の頻度分析手法に関する基礎的研究(Dissertation_全文)

AUTHOR(S):

寒川, 典昭

CITATION:

寒川, 典昭. 最大エントロピー法を用いた水文量の頻度分析手法に関する基礎的研究. 京都大学, 1989, 工学博士

ISSUE DATE:

1989-01-23

URL:

<https://doi.org/10.14989/doctor.r6765>

RIGHT:

最大エントロピー法を用いた水文量の
頻度分析手法に関する基礎的研究

1988年6月

寒 川 典 昭

正 誤 表

頁	行	誤	正
2	下11	Weighed	Weighted
7	9	Plinciple	Principle
20	下9	本説	本節
20	下3	Beta分布	Beta分布(=B(γ , δ))
43	4	1987年	1983年
47	下5	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots = \sigma_x^2$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots = \sigma_x^2$
53	2	であるから、式	であるから、式
134	2	図6.4	図6.4
134	5	図6.6	図6.6
147	6	7, 8 次モーメント	7, 8 モーメント
156	8	histogaram	histogram
157	5	前節	前項
157	下5	$\dots; \lambda_1^0, \lambda_2^0, \dots$	$\dots; \lambda_{1+}^0, \lambda_2^0, \dots$
159	下4	図8.6(f)	図8.4(f)
163	下5	M=30	M=28
170	表8.1	図8.6(f)	図8.4(f)
171	表8.2(a), (b), 表8.3	Anual	Annual
171	表8.3	guage	gauge
173	14 ~ 15	するとともにに	するとともに
174	10	条件付き最大エントロピー分布	条件付き分布

最大エントロピー法を用いた水文量の 頻度分析手法に関する基礎的研究

1988年6月

寒 川 典 昭

目 次

第1章 緒 論	1
参考文献	4
第2章 1 変数最大エントロピー分布	7
第1節 概 説	7
第2節 確率密度関数の推定	8
2.1 定式化	8
2.2 分布の導出	10
2.3 最大エントロピー分布の位置付け	13
2.4 パラメタ同定法	14
2.5 解の存在条件	16
2.6 モーメントの基準化	17
第3節 既存分布を母集団とした場合の適用性の検討	20
3.1 真のモーメントを用いた場合の適用性	20
3.1.1 仮定した母集団	20
3.1.2 適用性の検討	21
3.2 モーメントの安定性と感度分析	23
3.2.1 モーメントの安定性	23
3.2.2 モーメントの変動による感度分析	25
第4節 実データへの適用	26
4.1 実データの種類と抽出法	26
4.1.1 降水データ	27
4.1.2 流量データ	27
4.2 適用結果と考察	28
4.2.1 パラメタの値	28
4.2.2 ヒストグラムと推定分布	38
第5節 要 約	42
参考文献	43

第3章 2変数最大エントロピー分布	44
第1節 概説	44
第2節 確率密度関数の推定	45
2.1 定式化	45
2.2 分布の導出	46
2.3 パラメタ同定法	49
2.4 モーメントの基準化	52
第3節 既存分布を母集団とした場合の適用性の検討	54
3.1 真のモーメントを用いた場合の適用性	54
3.1.1 仮定した母集団	54
3.1.2 等しい形状母数をもつ母集団の場合	56
3.1.3 異なる形状母数をもつ母集団の場合	58
3.2 モーメントの安定性と感度分析	60
3.2.1 モーメントの安定性	60
3.2.2 データ数の増加と母集団推定	60
3.2.3 同数の異種データによる母集団推定	62
3.2.4 モーメントの変動による感度分析	62
第4節 実データへの適用	64
4.1 2地点流量の場合	64
4.2 年最大日降雨と対応日流量の場合	66
第5節 要約	68
Appendix A 2変数最大エントロピー分布からの2変数正規分布の導出	68
B 2変数最大エントロピー分布からの2変数指数分布（相関が零の場合）の導出	72
参考文献	74
第4章 多変数最大エントロピー分布	75
第1節 概説	75
第2節 確率密度関数の推定	76
2.1 定式化	76
2.2 分布の導出	77
2.3 パラメタ同定法	78
2.4 モーメントの基準化	82

第3節 既存分布を母集団とした場合の適用性の検討	84
3.1 仮定した母集団	84
3.2 適用性の検討	85
第4節 実データへの適用	90
4.1 3地点流量の場合	90
4.2 3地点8月降水の場合	91
第5節 要約	97
参考文献	97
第5章 条件付き最大エントロピー分布	99
第1節 概説	99
第2節 確率密度関数の推定	100
2.1 定式化	100
2.2 分布の導出	102
第3節 実データへの適用	104
3.1 年降水量の場合	105
3.1.1 モーメントの取り方による分布形の変化	105
3.1.2 条件の取り方による分布形の変化	110
3.2 年最大日降水量の場合	112
3.2.1 同時発生を仮定した年最大日降水量の推定	114
3.2.2 模擬降水による年最大日降水量の推定	114
第4節 要約	118
参考文献	119
第6章 情報の与え方と最大エントロピー分布	120
第1節 概説	120
第2節 任意関数の期待値を情報とした分布のパラメタ同定法	120
第3節 既存分布に対する適用性	122
3.1 単一関数の期待値を情報とした場合	122
3.2 異種関数の期待値を情報とした場合	126
3.3 同数の異種データによる母集団推定	131
第4節 実データへの適用	134
4.1 ヒストグラムに対する適用性	134

4.2 リターン・ピリオドの安定性	136
第5節 要 約	139
参考文献	140
第7章 最大エントロピー分布の諸特性の検討	141
第1節 概 説	141
第2節 諸特性の検討	141
2.1 原点にパルスをもつ分布	141
2.1.1 定式化	142
2.1.2 実データへの適用	143
2.2 高次モーメント適合	146
2.3 複数の極値をもつ分布への対応	147
2.3.1 理論的検討	147
2.3.2 ヒストグラムへの適用	149
第3節 要 約	150
Appendix	151
Appendix A 推定された密度関数のエントロピー最大性	151
B パラメタの一意性	151
参考文献	153
第8章 最大エントロピーヒストグラム	154
第1節 概 説	154
第2節 c-ヒストグラムの推定	155
2.1 定式化	155
2.2 パラメタ同定法	157
第3節 実データへの適用	159
第4節 要 約	172
参考文献	172
第9章 結 論	173
謝 辞	176

第 1 章 結 論

確率分布を用いて 1 変数から多変数までの水文量の母集団分布を推定するという本研究の課題は、水文統計学において中心的役割をなすものである。

一般に、水工計画では、降水量・河川流量等の水文データを用いて、対象とする大きさの確率水文量のリターン・ピリオドを計算したり、逆にリターン・ピリオドを指定して、それに応じた確率水文量を算定することが要求される。これらを計算するためには、対象とする水文量の母集団を推定しなければならない。

1 変数統計におけるその重要性は、河川計画を始め多くの場面で用いられていることを見れば、論を待つまでもなからう。

ところが、水工計画が 2 変数以上の多変数確率分布を用いて議論されなければならない問題は非常に多い。例えば長沢¹⁾が述べているように、内水排除計画の場合、洪水の継続時間とその前後に降る雨（内水降雨）が主として対象とされる水文量であるが、さらに経済効果を把握するためにはピーク水位もまた重要な要素となり、これらの水文量の関連を統計的に検討することが要請される。また長尾²⁾の指摘にみられるように、河川堤防による治水計画では、洪水時の、ピーク水位、流量、流速等が破堤・溢水に関係する主たる因子となり、それらの諸量を確率変数とみなした統計解析が必要とされる。石原・長尾³⁾が淀川をモデルケースとして取り扱った、2 本の支川が合流して 1 本の本川となり、3 川の堤内地が防災の対象となる場合も、基本的には 3 変数分布が利用される好例であろう。このように、計画の対象となる複数の水文量の間に相関が存在する場合、変数の数にみあった多変数確率分布を用いることは、合理的な水工計画を作成する上で、極めて重要なことである。

また、我々が取り扱う水文量は、統計解析する上で十分に整備されていることはむしろ稀で、水文データの推定、および模擬発生の必要性が生じることがある。それは、特定の任意個数の変数を指定したとき、対象とする他の任意個数の変数の統計解析をしなければならない場合であり、このとき条件付き確率分布が要求される。正規分布の回帰直線の利用、長尾・角屋⁴⁾による基準地点降雨に基づいた対象地点降雨の模擬発生等は、水文統計において条件付き分布の必要性を示す良い具体例と言えよう。

さて、確率分布による水文統計の従来の研究は、大別すると、確率分布の種類の問題とパラメタ推定の問題になると考えられる。

確率分布の種類の問題は、さらに取り扱う確率変数の数により、1 変数、2 変数、多（任意個数の）変数分布に分類される。また、2 変数以上の多変数分布から条件付き分布が定義され

ている。

1 変数分布は、一般に、次のように分けられる。

(i) 正規分布を前提とした分布：正規分布，対数正規分布。

(ii) Pearson III 型分布（Gamma 分布）を前提とした分布：Pearson III 型分布，対数 Pearson III 型分布，指数分布。

(iii) 極値分布：Gumbel 分布，Weibull 分布，平方根指数型最大値分布⁵⁾等。

(i)，(ii) はさらに、対数変換の有無，母数の数，母数の値により，(iii) は母集団分布の種類により，個々の分布に対応付けられている。

2 変数分布になると，その数は急減する。従来，水文統計に用いられてきた分布としては，2 変数正規分布⁶⁾，2 変数 Gamma 分布（特殊な場合として 2 変数指数分布を含む）^{7), 8)}，および Freund の 2 変数指数分布⁹⁾に限られよう。

これが 3 変数以上の多変数分布では，実用的に用いられるようにその特性が明らかにされているのは多変数正規分布のみである。数学的興味においては，Krishnamoorthy・Parthasarathy¹⁰⁾により多変数 Gamma 分布が提案されているが，水文統計に導入されるには解決されるべき問題が少なくない。

条件付き分布では，上述の 3 つの 2 変数分布から，1 つの変数を条件とし，残りの 1 つの変数を対象とした，条件付き正規分布⁶⁾，条件付き Gamma 分布^{7), 11)}，Freund の条件付き指数分布⁹⁾が知られている。しかし，条件とする変数の数も，対象とする変数の数も制限しない条件付き分布については，著者の知るかぎり，今まで少なくとも水文統計の分野では，未開発のままとなっていた。

一方，確率分布のもつパラメタ推定法は，点推定法と区間推定法に大別され，点推定法の中には，図式推定法，最小 2 乗法，モーメント法，最尤法がある。また，最近 Greenwood ら¹²⁾により PWM (Probability Weighed Moments) 法が提案され，日本では，竹内ら¹³⁾により，その有効性が検討されつつあるが，まだ緒についたばかりと言えよう。さらに，少し趣を異にするが，水文統計では小標本からの統計解析を余儀なくされていることから，パラメタ推定においてぬぐいされない不確定さを評価しようとした寒川ら¹⁴⁾の研究もある。

こうしてみると，確率分布による水文統計の社会的要請は，1 変数の場合は言うに及ばず，2 変数，3 変数以上の多変数，および条件付きの場合においても，非常に大きいことがわかる。ところが，それに対処する理論面での研究は，上でも触れたように，特に 2 変数以上の多変数や条件付き分布において，まだまだ不十分である。さらに，水文量は，一般に，原型のままで正規分布に従うことは稀であり，非正規性をもつ分布が要求される。

このような状況から，水文統計の分野において，水文量の宿命とも言える小標本性をうまく取り込み，理論的背景を持ち，利用において簡便で，水文データに対する全体的な適合度がよ

く、さらに対象とする推定値（例えば確率水文量）が妥当で、安定した値を与え、しかも変数の数に制限をもたない確率分布の開発が望まれていた。

こうした水文統計の要請を可能にする確率分布を導出するための基礎概念に、Shannon¹⁵⁾のエントロピーがある。高棹は早くからこの概念の水文現象への適用を指摘し、1969年、高棹・池淵¹⁶⁾はBlack-Boxとして取り扱われてきた長期流出系の内部構造を確率論的に探るためにエントロピーの概念を用いた。その後、高棹・池淵・寒川^{17), 18)}によって、この研究は長期流出系に獲得情報量最大仮説を置くモデルへと発展させられ、貯留状態の遷移確率構造が議論された。また、前者の研究に端を発して、高棹・池淵は、降水観測がもたらす情報量¹⁹⁾、水文系列の分布特性²⁰⁾および降雨・流出変換過程²⁰⁾をエントロピー的に解釈した。この水文系列が従う分布のエントロピー的解釈は、水文量の母集団分布が最大エントロピー分布によって十分よく再現されることを示唆していた。

本研究は、このような確率分布による水文統計の必要性和確率分布の開発状況を踏まえながら、上述の条件を満足せしめる分布として、高棹・池淵によってその有効性が暗示されたShannonのエントロピーから導かれる最大エントロピー分布を取り上げ、水文統計のためのすぐれた確率分布を目差して、この分布を拡張・発展させたものである。すなわち、本論文は、水文量が小標本であることを前提として、1変数最大エントロピー分布の水文統計への導入と改良、および、2変数、多変数、条件付きへの拡張について論じたものであり、その内容を簡単に述べれば以下のようなものである。

ただし、最大エントロピー分布のパラメタは、前述のパラメタ推定法のどれをも用いるのではなく、情報の与えかた、言い換えれば、制約条件を満足するように決定されることになっている。

第2章では、第3章以後の理論展開の基礎となる1変数最大エントロピー分布の理論を要約して提示する。特に、水文統計への最大エントロピー分布の導入の妥当性、および既存分布の導出を踏まえた最大エントロピー分布の分布系としての位置付けの議論に力点をおいている。さらに、既存分布を母集団と仮定した場合と、多くの時間単位の降雨・流量データからなる実データを用いた場合の両者から、統計モーメントを情報とした最大エントロピー分布の適用性を検討する。

第3章では、多変数統計の基礎として、2変数最大エントロピー分布を提案する。ここでも、前章に準じて、2変数Gamma分布を母集団と仮定した場合と、相関のある2地点で観測された実データを用いた場合の両者から、統計モーメントを情報とした2変数最大エントロピー分布の適用性を検討する。また、Appendixとして、2変数最大エントロピー分布からの、2変数正規分布と相関のない2変数指数分布の導出過程を示す。

第4章では、最も一般的な場合の確率分布として、前章の2変数を多変数に拡張した最大エ

ントロピー分布を提案する。多変数とは、3変数以上の有限個数の変数を指している。ここでも、前2章に従って、3変数正規分布を母集団と仮定した場合と、相関のある3地点で観測された実データを用いた場合の両者から、統計モーメントを情報とした最大エントロピー分布の適用性を検討する。

第5章では、前章の多変数最大エントロピー分布から、条件付き分布の定義に従って、条件とする変数の個数にも、対象とする変数の個数にも制限を加えない条件付き最大エントロピー分布を導出する。また、多地点で観測された相関のある実データを用いて、この分布の特徴と有用性を把握する。

第6章では、前章までの統計モーメントで情報を与える方法をさらに発展させ、対象とする母集団に対して、パラメタ数が少なくて適合度のよい情報の与え方と、最大エントロピー分布の関係を論じる。まず、1変数 Gamma分布を母集団として、適切な情報の与え方を模索し、そこで得られた候補となる分布を年最大日降水量に適用して、この水文量に対して適切な情報の与え方を提案する。

第7章では、前章までに明らかにされなかった、最大エントロピー分布の諸特性を検討する。具体的には、原点にパルスを持つ最大エントロピー分布とその適用性、高次モーメントまで制約条件とした最大エントロピー分布の適用性、および複数の極値をもつヒストグラムへの最大エントロピー分布の対応、の3項目である。また、Appendixに、パラメタ同定の基礎知識として必要な、2つの証明を紹介しておく。

第8章では、滑らかな曲線を有する最大エントロピー分布が存在しない場合に対処するために、第2章の最大エントロピー分布の代用として、最大エントロピーヒストグラムの水文統計への導入を考察する。ここでは、理論式の要約とともに、このヒストグラムの実データへの適用をはかり、最大エントロピー分布との比較において、このヒストグラムの特徴を論ずる。

第9章は研究の結論である。

なお、ここで取り扱っている確率変数はすべて連続的な場合であることを付記しておく。

参 考 文 献

- 1) 長沢敏夫：複合確率と河川工事計画における応用について，建設省直轄工事技術報告，第14回，pp.791～796，1960年．
- 2) 長尾正志：水文統計における多変数確率分布理論——二変数統計を中心として——，水工学に関する夏期研修会講義集（Aコース），土木学会水理委員会，pp.A-4-1～A-4-19，1975年，8月．

- 3) 石原安雄・長尾正志：計画高水流量の二変数確率論的研究，京都大学防災研究所年報，第6号，pp.181～188，1963年7月。
- 4) 長尾正志・角屋睦：二変数ガンマ分布とその適用に関する研究（4）——短時間降雨の模擬発生——，京都大学防災研究所年報，第14号，B，pp.77～85，1971年4月。
- 5) 江藤剛治・室田 明・米谷恒春・木下武雄：大雨の頻度，土木学会論文集，第369号/II-5，pp.165～174，1986年5月。
- 6) 例えば，長尾正志：水文統計解析，河川工学百年の歩みと淀川，p.88，1978年9月。
- 7) 井沢竜夫：二変数 Γ -分布について（降水量の分布 第2報），気象と統計，第4巻，第1号，pp.9～15，1953年。
- 8) 井沢竜夫：二変数 Γ -分布について（続），気象と統計，第4巻，第2号，pp.15～19，1953年。
- 9) Freund, J. E. : A bivariate extension of the exponential distribution, American Statistical Association Journal, Vol.56, pp.971～977, Dec., 1961.
- 10) Krishnamoorthy, A. S., and Parthasarathy, M. : A Multivariate gamma-type distribution, The annals of mathematical statistics, Vol.22, pp.549～557, 1951.
- 11) 長尾正志：水文統計解析，河川工学百年の歩みと淀川，p.89，1978年9月。
- 12) Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R. : Probability weighted moments : Definition and relation to parameters of several distributions expressible in inverse form, Water Resour. Res., Vol.15, No.5, pp.1049～1054, Oct., 1979.
- 13) 竹内邦良・土屋一仁：正規分布のパラメータのPWM解，第31回水理講演会論文集，pp.191～196，1987年2月。
- 14) 寒川典昭・荒木正夫・渡辺輝彦：確率分布の推定母数の不確定性評価法，土木学会論文集/II-6，pp.133～141，1986年11月。
- 15) 例えば，笠原芳郎：情報理論と通信方式，共立出版，p.150，1965年10月。
- 16) 高棹琢馬・池淵周一：長期間流出機構の情報理論的研究，京都大学防災研究所年報，第12号，B，pp.273～293，1969年3月。
- 17) 高棹琢馬・池淵周一・寒川典昭：エントロピー・モデルに関する2，3の考察，京都大学防災研究所年報，第24号，B-2，pp.143～158，1981年4月。
- 18) 高棹琢馬・池淵周一・寒川典昭：長期流出系のエントロピーモデル，土木学会論文報告集，第337号，pp.111～117，1983年9月。
- 19) 高棹琢馬・池淵周一：降水観測がもたらす情報量とその観測網配置計画への応用，土木学会論文報告集，第234号，pp.83～95，1975年2月。

- 20) 高棹琢馬・池淵周一：エントロピー的にみた降雨・流出変換特性とそのモデル化，京都大学防災研究所年報，第23号，B-2，pp.193～209，1980年4月．

第2章 1変数最大エントロピー分布

第1節 概 説

対象とする母集団にいかなる分布形をあてはめるかという問題に対して、思想的な妥当性を与えた分布系の1つが、最大エントロピー分布である。本論文は、水文統計を目的として、この最大エントロピー分布について詳細に議論したものである。

本章では、1変数最大エントロピー分布について考察する。^{1), 2)} Sonuga³⁾は、水文量はしばしば小標本であることに着目して、2次モーメントまでを制約条件とした1変数最大エントロピー分布による水文量の頻度分析法を提案した。Sonugaの研究は、最大エントロピー原理(Plinciple of Maximum Entropy)を用いることにより、分布形の採用に思想的な妥当性を与えた点で高く評価されるが、パラメタ同定法が複雑なため、3次以上の統計モーメント、および他の関数を制約条件に加え難い(情報として取り入れ難い)という欠点を有していた。ところが、情報理論の分野で、Wragg・Dowson⁴⁾が提案したパラメタ同定法を用いると、任意次数までの統計モーメントを情報とした最大エントロピー分布のパラメタは容易に求まる。

したがって、第2節では、任意関数を情報とした1変数最大エントロピー分布を説明し、パラメタ同定法では、統計モーメントを情報とした場合に絞って、Wragg・Dowsonの方法を採用する。また、この理論式の説明にあたり、最大エントロピー分布による水文量の頻度分析の妥当性を、その分布のもつ特性と水文量の小標本性より解き明かす。一方、高棹・池淵⁵⁾は最大エントロピー分布から、水文量の頻度分析に用いられるいくつかの既存分布が導出されることと、水文量の従う分布形が時間単位の取り方により変動することとの関連を、対象とする水文量のどの特性が安定しているかという立場から論じている。本節では、この研究を受けて、最大エントロピー分布から多くの既存分布を導出できることをレビューしながら、この分布の分布系としての位置付けを明確にする。さらに、Einbu⁶⁾によって導かれた、モーメントで情報を与えた1変数最大エントロピー分布の存在条件とWragg・Dowson⁴⁾が行っているモーメントの基準化を提示しておく。前者は、パラメタ同定が可能かどうかの情報となり、後者は、パラメタ同定のときの計算上の利点となる。なお、モーメントの基準化は、第3章、第4章でも行っているが、上述と同じ理由のためであるので、その章の概説と要約では特に触れない。

第3節では、仮定した母集団に対する適用性を、真のモーメントを使った場合について検討し、さらに、モーメントの安定性を把握するとともに、モーメントの変動に伴う推定分布の感度を分析する。

第4節では、種々の時間単位の降水データと流量データを取り上げ、推定分布のパラメタの値、およびデータのヒストグラムに対する適合性を検討する。

なお、頻度解析に用いる水文データは、定常・独立性、均質性が要求されよう。前者は、水文量時系列に周期成分、トレンド成分、従属成分等が含まれていないこと、後者は水文量が同一の母集団から生起したとみなせること、例えば発生原因の異なる降雨の分離の必要性の検討等である。本研究では、このような問題に対して、注意しながらデータを調べているが、前者については、水文量時系列のグラフ化からみて、特に取り徐くべき大きな成分が見られなかったこと、後者については、分離すればデータ数が減少するという問題が発生すること、等の諸理由からデータは加工せず生のまま用いていることを付記しておく。

第2節 確率密度関数の推定

2.1 定式化

確率変数 x の確率密度関数を $p(x)$ とすると、エントロピーは次式で与えられる。¹⁾

$$H(x) = - \int p(x) \ln p(x) dx \quad (2.1)$$

一方、確率密度関数が具備すべき条件と任意関数 $g_r(\cdot)$ の期待値は、以下のように表現される。

$$\int p(x) dx = 1 \quad (2.2)$$

$$\int g_r(x) p(x) dx = E[g_r(x)], r=1, 2, \dots, N \quad (2.3)$$

ここに、 $E[\cdot]$ は期待値記号である。

さて、本節 2.2 で記述するように、 x の定義域を有限区間とした場合、式 (2.2) を制約条件として式 (2.1) を最大にする $p(\cdot)$ は、その区間内で一様分布となる。すなわち、有限区間内で定義された確率変数 x が、その生起に関して情報が皆無であるとき、最大エントロピー分布はどの x に関しても等しい確率密度を与える。

従って、式 (2.2)、(2.3) を制約条件として、式 (2.1) を最大にする $p(\cdot)$ を求めること (最大エントロピー法) は、母集団から得られる情報は $g_r(\cdot)$ の期待値の形で取り入れ、それ

以外はできるだけ確率変数がランダムに発生する分布を算定していることになる。換言すれば、それは情報 ($g_r(\cdot)$) の与え方には主観が入るが、それ以外は最も偏見の少ない、最も客観的な分布を求めていることになる。筆者は、水文量のようにデータ数が少なく不確定性の強い自然現象の説明に、このような確率分布を導入することはきわめて妥当な立場と考える。

又、従来の水文量の統計解析は、多くの確率分布から対象とする水文量によく適合する分布が選ばれ、その分布の母数をできるだけ真値に近い値に決定するために積率法、最尤法等の推定法のなかから最適な手法が選択される2段階決定法であった。これに対し、最大エントロピー法は、母集団の分布をあらかじめ仮定せず、あるいは経験によって決定せず、分布形は $g_r(\cdot)$ にのみ依存して決まり、その母数は式 (2.2), (2.3) を満足するように決定される1段階決定法であるといえる。したがって、意志決定者は対象とする水文量の特性を十分に説明し得る $g_r(\cdot)$ のみを選択すればよいことになるため、最大エントロピー法は一貫した概念での統計解析を可能ならしめよう。

そこで、上記の問題をラグランジュの未定乗数法で解く。ラグランジュ関数は

$$L=H(x)+(\lambda_0-1)\{1-\int p(x)dx\}+\sum_{r=1}^N \lambda_r \{E[g_r(x)]-\int g_r(x)p(x)dx\} \quad (2.4)$$

となる。ここに、 $\lambda_r (r=0,1,\dots,N)$ はラグランジュ乗数である。上式を $p(\cdot)$ について変分をとり零とおくと

$$-\{1+\ln p(x)\}-(\lambda_0-1)-\sum_{r=1}^N \lambda_r g_r(x)=0 \quad (2.5)$$

となり、上式より $p(\cdot)$ の最大エントロピー推定値

$$p(x)=\exp\{-\lambda_0-\sum_{r=1}^N \lambda_r g_r(x)\} \quad (2.6)$$

が得られる。ここに、ラグランジュ乗数 λ_r は推定分布のパラメタ (母数) であることがわかる。ただし、上式中の λ_0 は式 (2.2) を満足するように $\lambda_r (r=1,2,\dots,N)$ に依存して決定されるため、他の確率分布との対比で式 (2.6) を論じるとき、パラメタ数は N 個と呼ぶべきである。

式 (2.6) は一般的な1変数最大エントロピー分布であり、種々の $g_r(\cdot)$ を与えることができ

る。

2.2 分布の導出

最大エントロピー分布は情報の与え方により，形式的には無限個の分布形を記述することができる。その中には統計学においてよく知られている多くの分布が存在する。ここでは，主要な既存分布を1変数最大エントロピー分布から導出し，2.3を論述するための基礎知識とする。

1) 1変数一様分布

確率変数の存在範囲を $[L_1, L_2]$ として式(2.2)を次のように書き換える（式(2.3)は用いない）。

$$\int_{L_1}^{L_2} p(x) dx = 1 \quad , \quad L_2 > L_1 \quad (2.7)$$

このとき，式(2.6)は

$$p(x) = \frac{1}{L_2 - L_1} \quad (2.8)$$

となり，これは1変数一様分布である。

2) 1変数正規分布⁸⁾

$$g_1(x) = (x - \mu)^2 \quad (2.9)$$

とし，確率変数の存在範囲を $(-\infty, \infty)$ として式(2.2)，(2.3)を次のように書き換える。

$$\int_{-\infty}^{\infty} p(x) dx = 1 \quad (2.10)$$

$$\int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx = \sigma^2 \quad (2.11)$$

ここに， μ ， σ^2 は定数である。

このとき，式(2.6) は

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \quad (2.12)$$

となり，これは 1 変数正規分布である。

3) 1 変数対数正規分布⁹⁾

$$g_1(x) = (\ln x - \lambda)^2 \quad (2.13)$$

とし，確率変数の存在範囲を $[0, \infty)$ として式(2.2)，(2.3)を次のように書き換える。

$$\int_0^{\infty} p(x) dx = 1 \quad (2.14)$$

$$\int_0^{\infty} (\ln x - \lambda)^2 p(x) dx = \zeta^2 \quad (2.15)$$

ここに， λ ， ζ^2 は定数である。

このとき，式(2.6) は

$$p(x) = \frac{1}{\sqrt{2\pi\zeta^2} x} \exp \left\{ -\frac{(\ln x - \lambda)^2}{2\zeta^2} \right\} \quad (2.16)$$

となり，これは 1 変数対数正規分布である。

4) 1 変数指数分布¹⁰⁾

$$g_1(x) = x \quad (2.17)$$

とし，確率変数の存在範囲を $[0, \infty)$ として式(2.2)，(2.3)を次のように書き換える。

$$\int_0^{\infty} p(x) dx = 1 \quad (2.18)$$

$$\int_0^{\infty} xp(x)dx = \eta \quad (2.19)$$

ここに， η は定数である。

このとき，式(2.6) は

$$p(x) = \frac{1}{\eta} \exp\left(-\frac{x}{\eta}\right) \quad (2.20)$$

となり，これは 1 変数指数分布である。

5) Gumbel 分布¹¹⁾

$$g_1(x) = \frac{x - \alpha}{\beta} \quad (2.21)$$

$$g_2(x) = \exp\left\{-\left(\frac{x - \alpha}{\beta}\right)\right\} \quad (2.22)$$

とし，確率変数の存在範囲を $(-\infty, \infty)$ として式(2.2)，(2.3)を次のように書き換える。

$$\int_{-\infty}^{\infty} p(x)dx = 1 \quad (2.23)$$

$$\int_{-\infty}^{\infty} \left(\frac{x - \alpha}{\beta}\right) p(x)dx = \gamma \quad (2.24)$$

$$\int_{-\infty}^{\infty} \exp\left\{-\left(\frac{x - \alpha}{\beta}\right)\right\} p(x)dx = 1 \quad (2.25)$$

ここに， α ， β は定数， γ は Euler 定数である。

このとき，式(2.6) は

$$p(x) = \frac{1}{\beta} \exp\left[-\left(\frac{x - \alpha}{\beta}\right) - \exp\left\{-\left(\frac{x - \alpha}{\beta}\right)\right\}\right] \quad (2.26)$$

となり，これはGumbel分布である。

2.3 最大エントロピー分布の位置付け

Pearson は，統計学に出てくる確率密度関数の多くは，微分方程式

$$\frac{1}{p(x)} \frac{dp(x)}{dx} = \frac{k_0 + x}{k_1 + k_2 x + k_3 x^2} \quad (2.27)$$

の解であることを示した。ここに， k_0, k_1, k_2, k_3 は定数である。 $p(x)$ は定数相互間の関係により，3個のMain types(Pearson I型，IV型，VI型)と10個のTransition types (Pearson II型，III型，V型，VII型，VIII型，IX型，X型，XI型，XII型，Normal curve)に分類され，それらはPearson系の分布と呼ばれている。¹²⁾ その中には，J型分布，U型分布，対称分布，非対称分布があり，また両側有限範囲の分布から片側有限範囲・片側無限範囲，および両側無限範囲に至るまで多様の分布曲線が存在する。

いま，最大エントロピー分布とPearson系の分布の比較をすると，最大エントロピー分布の制約条件，すなわち式(2.3)は，Pearson系の分布の定数相互間の関係に，任意関数 $g_r(\cdot)$ を制約条件とした最大エントロピー分布，すなわち式(2.6)は，式(2.27)の微分方程式に相当する。

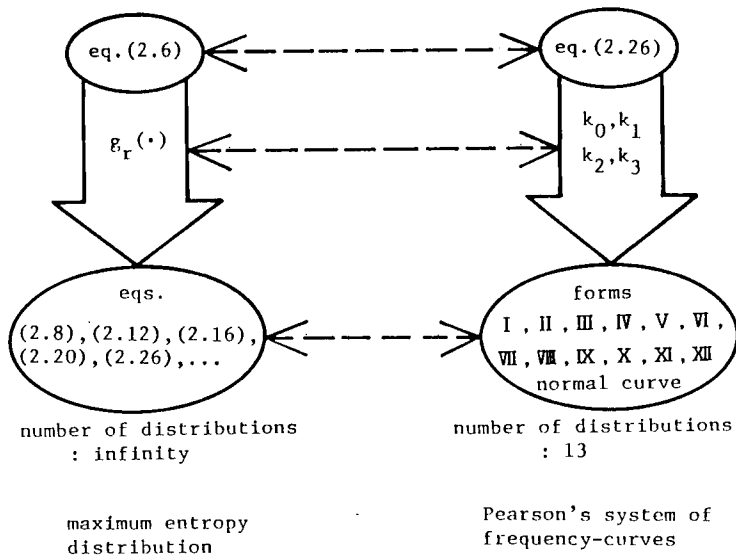


図2.1 最大エントロピー分布とPearson系の分布の比較

このことは、最大エントロピー分布は Pearson系の分布等と同レベルで議論されるべきものであり、具体的な $g_r(\cdot)$ を与えて得られた最大エントロピー分布、たとえば式(2.26)は、Pearson 系の分布の 1 つ、たとえば PearsonⅢ型分布と対応されるべきものであることを示している。図2.1 はそれを模式的に表現したものである。ただし、最大エントロピー分布は無限個数の分布を有することより、Pearson系の分布よりさらに汎用性に富んでいる。

一方、1 つの関数から導かれる関数を項とする無限級数によって構成される分布系として Gram-Charlier の級数がある。¹³⁾ これは、確率密度関数が 2 つの一般形で表現されていること、連続分布と離散分布の両者を有していること、の理由により上述の対比から除外したが、広い意味において上記 2 分布形と同等な取り扱いがなされ得るものである。

なお、ここでの議論は、Pearson系の分布が 1 変数であるため 1 変数最大エントロピー分布に限ったが、最大エントロピー分布は当然 2 変数以上にも容易に議論を拡張し得ることを付記する。

2.4 パラメタ同定法

いま、式(2.3)として、取り扱いが容易で 4 次までは分布の形状との関係が明らかにされている原点まわりの統計モーメントを採用し、積分区間は水文量を対象としていることから $[0, \infty)$ とする。そのとき、式(2.2)、(2.3)は次式のように書き換えられる。

$$\int_0^{\infty} p(x) dx = 1 \quad (2.28)$$

$$\int_0^{\infty} x^r p(x) dx = \mu_r, \quad r=1, 2, \dots, N \quad (2.29)$$

ここに、 μ_r は原点のまわりの r 次のモーメントである。上 2 式を制約条件とした 1 変数最大エントロピー分布は、式(2.6) より次式となる。

$$p(x) = \exp(-\lambda_0 - \sum_{r=1}^N \lambda_r x^r) \quad (2.30)$$

上式を式(2.29)に代入して、式(2.28)を用いて整理すると

$$\int_0^{\infty} x^r \exp\left(-\sum_{r=1}^N \lambda_r x^r\right) dx = \mu_r \int_0^{\infty} \exp\left(-\sum_{r=1}^N \lambda_r x^r\right) dx, \quad r=1,2,\dots,N \quad (2.31)$$

なる方程式が得られる。したがって、式(2.30)のパラメタ $\lambda_r (r=1,2,\dots,N)$ は式(2.31)を解くことにより決定され、 λ_0 は式(2.30)を式(2.28)に代入して得られた次式から算定される。

$$\lambda_0 = \ln\left\{\int_0^{\infty} \exp\left(-\sum_{r=1}^N \lambda_r x^r\right) dx\right\} \quad (2.32)$$

ここで、式(2.31)の非線形方程式を λ_r に関して解くために、収束過程が明示でき、その収束度合をみながら補正量（残差）を修正することが可能であり、また以下の計算では十分な精度の解を得ることができた Wragg・Dowson⁴⁾ の方法を採用する。

いま、 λ_r の近似値として α_r を与え、残差を ϵ_r とおくと、次式が成立する。

$$\lambda_r = \alpha_r + \epsilon_r \quad (2.33)$$

そこで、上式を式(2.31)に代入して、 ϵ_r に関して Taylor 展開し、 $(\epsilon_r)^2$ 以上の微小量の項を無視して整理すると

$$\sum_{j=1}^N (A_i + j - \mu_j A_j) \epsilon_j = A_i - \mu_i A_0, \quad i=1,2,\dots,N \quad (2.34)$$

が得られる。ここに、

$$A_i = \int_0^{\infty} x^i \exp\left(-\sum_{r=1}^N \alpha_r x^r\right) dx, \quad i=1,2,\dots \quad (2.35)$$

である。

式(2.34)は ϵ_j に関する N 次元 1 次連立方程式であり、変数の数と方程式の数が一致するので解き得る。

実際、式(2.30)のパラメタは、式(2.34)を繰り返し解くことにより求められる。具体的には次のように計算する。

- 1) 初期値 $\alpha_j=0.0$ と仮定する。経験的に言えば、この場合が最も収束性がよい。

- 2) 式(2.34)の連立方程式を ϵ_j に関して解く。
- 3) 得られた解を式(2.33)に代入して λ_j を計算し、これを次の α_j とする。
- 4) ϵ_j が十分小さくなるまで(10^{-8} 程度)，2)，3)を繰り返す。
- 5) ϵ_j が微小となったとき， α_j を推定パラメタとみなす。
- 6) 得られた推定パラメタを式(2.32)に代入して λ_0 を算定する。

$N=4$ の場合， $\epsilon_j \leq 10^{-8}$ ($j=1,2,3,4$) で収束したとみなすと，モーメントの与え方によって変動するが，繰り返し回数は以後に取り上げた例で5 ～40回程度である。

2.5 解の存在条件

式(2.30)で与えられる1変数最大エントロピー分布は，常に存在するとはかぎらない。すなわち，最初の $(N-1)$ 個のモーメント $\{\mu_1, \mu_2, \dots, \mu_{N-1}\}$ に対する1変数最大エントロピー分が存在して，そのパラメタが $\{\lambda_1, \lambda_2, \dots, \lambda_{N-1}\}$ で与えられれば， μ_N が次の2つの条件を満たすときのみ， $\{\mu_1, \mu_2, \dots, \mu_N\}$ に関する1変数最大エントロピー分布が存在する。⁶⁾

$$\mu_N \leq \mu_{N, \max} = \int_0^\infty x^N \exp\left(-\sum_{i=0}^{N-1} \lambda_i x^i\right) dx \quad (2.36)$$

$$\mu_N = \mu_{2i} > \mu_{2i, \min} = \begin{vmatrix} 1 & \mu_1 & \cdots & \mu_{i-1} & \mu_i \\ \mu_1 & \mu_2 & \cdots & \mu_i & \mu_{i+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \mu_i & \mu_{i+1} & \cdots & \mu_{2i-1} & 0 \end{vmatrix} \cdot \begin{vmatrix} 1 & \mu_1 & \cdots & \mu_{i-1} \\ \mu_1 & \mu_2 & \cdots & \mu_i \\ \vdots & \vdots & & \vdots \\ \mu_{i-1} & \mu_i & \cdots & \mu_{2(i-1)} \end{vmatrix}^{-1} \quad (2.37)$$

ここに，式(2.36)は $\{\lambda_0, \lambda_1, \dots, \lambda_{N-1}\}$ が既知のときにのみ，式(2.37)は N が偶数のときにのみ計算されるということに注意を要する。

いま， $N=2$ を考えると，式(2.36)，(2.37)より

$$\mu_{2, \max} = 2(\mu_1)^2 \quad (2.38)$$

$$\mu_{2, \min} = (\mu_1)^2 \quad (2.39)$$

となり，1変数最大エントロピー分布は

$$(\mu_1)^2 < \mu_2 \leq 2(\mu_1)^2 \quad (2.40)$$

のときに存在する。分散を σ^2 と書くと上式は次のように書き換えられる。

$$0 < \sigma / \mu_1 \leq 1 \quad (2.41)$$

Sonuga³⁾の提案した2次モーメントまでを用いた1変数最大エントロピー分布は，上式の条件下で議論されていることになる。

2.6 モーメントの基準化

式(2.29)における右辺の値，すなわちデータから求められる統計モーメントは，高次になると大きな値となって取り扱いに困難を生じる。そこで，Wragg・Dowson⁴⁾による確率変数の平均が1となるような変数変換

$$u = x / \mu_1 \quad (2.42)$$

を行うと， u に関する確率密度関数 $q(u)$ は，Jacobian $J(\cdot)$ を介して

$$q(u) = J\left(\frac{x}{u}\right)p(x) \quad (2.43)$$

となる。ここに

$$J\left(\frac{x}{u}\right) = \mu_1 \quad (2.44)$$

であるから，式(2.43)は

$$q(u) = p(x) \mu_1 \quad (2.45)$$

あるいは

$$p(x) = q(u) / \mu_1 \quad (2.46)$$

と表現できる。

上述の変数変換を行うと，式(2.28)，(2.29)は次のように書き換えられる。

$$\int_0^\infty q(u) du = 1 \quad (2.47)$$

$$\int_0^\infty u^r q(u) du = \mu_r / (\mu_1)^r \equiv m_r, \quad r=1, 2, \dots, N \quad (2.48)$$

ここに， $m_0 = m_1 = 1$ であるため，式(2.48)で定義された m_r を基準化されたモーメントと呼ぶことにする。

次に， $p(\cdot)$ と $q(\cdot)$ のエントロピーの関係を求める。式(2.42)と式(2.46)を式(2.1)に代入すると

$$\begin{aligned} H(x) &= - \int_0^\infty p(x) \ln p(x) dx \\ &= - \int_0^\infty q(u) \ln q(u) du + \ln(\mu_1) \end{aligned} \quad (2.49)$$

$$\equiv H(u) + \ln(\mu_1) \quad (2.50)$$

が得られる。

上式は，エントロピー $H(x)$ が最大るとき，変数変換したエントロピー $H(u)$ も最大になることを示している。従って，基準化されたモーメントによる制約条件の下で $H(u)$ を最大にする問題は，2.1 で定式化した問題と等価である。実際の適用では $q(\cdot)$ を求めておき，しかる後に式(2.46)により $p(\cdot)$ に変換すればよい。

いま，式(2.48)で定義した m_r を用いると， $N=2$ の場合 m_2 のみに依存して1変数最大エントロピー分布が定まる。図2.2はその関係を図示したものであり， λ_0 と λ_2 は m_2 の増大に対して

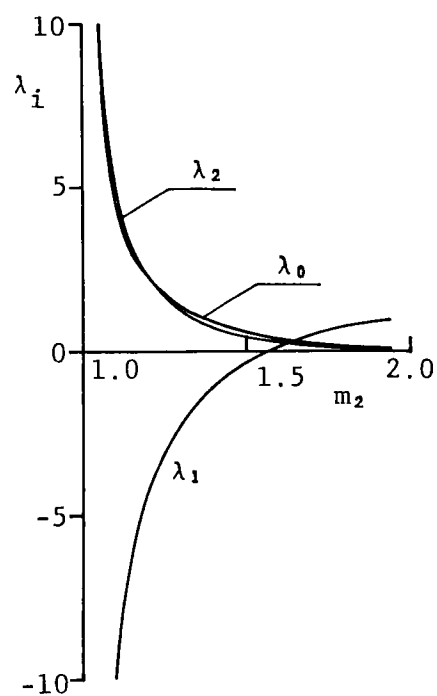


図2.2 m_2 と λ_0 , λ_1 および λ_2 の関係

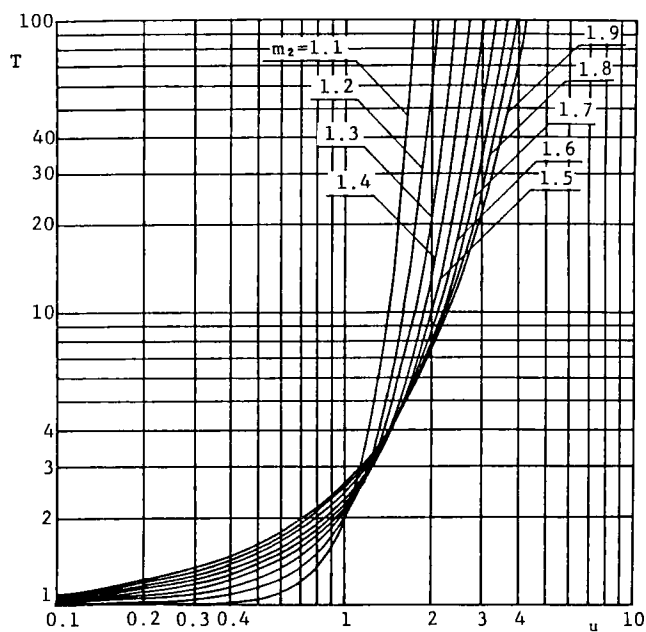


図2.3 m_2 をパラメタとした u と確率年との関係

指数関数的に減少し、 λ_1 は対数関数的に増加すること、および $m_2 < 1.5$ の区間ではパラメタが m_2 の微小変化に対して大きく変動することがわかる。また、 m_2 をパラメタとした u と確率年 ($=T$) との関係を図示したのが図2.3 であり、2 モーメント適合 (最初の N 個のモーメントを用いて1変数最大エントロピー分布を求めることを N モーメント適合と呼ぶ) の確率年は、この図から u を指定することにより、容易に読み取ることができる。

第3節 既存分布を母集団とした場合の適用性の検討

3.1 真のモーメントを用いた場合の適用性

3.1.1 仮定した母集団

水文量は1つの母集団から生じたデータであると仮定されるが、水文統計上、その母集団は未知であるため、データから得られたヒストグラムをうまく表現できる確率分布が導入される。しかし、この場合、導入した確率分布が、母集団をどの程度正確に記述しているかを、直接に測ることは難しい。そこで、本説では母集団を、次のGamma分布とBeta分布に仮定して、以下の議論を展開する。

パラメタ α, β をもつ Gamma分布 ($=G(\alpha, \beta)$) は、密度関数

$$p(x) = \frac{1}{\beta^{\alpha+1} \Gamma(\alpha+1)} x^{\alpha} \exp(-x/\beta) \quad , \alpha > 0, \beta > 0, x > 0 \quad (2.51)$$

より定義され、原点のまわりの r 次モーメントは

$$\mu_r = (\alpha+1)(\alpha+r-1) \cdots (\alpha+1) \beta^r \quad , \quad r=1,2,\cdots,N \quad (2.52)$$

で与えられる。同様に、パラメタ γ, δ をもつBeta分布では、式(2.51),(2.52) は次の2式で置き換えられる。

$$p(x) = \frac{(\gamma + \delta + 1)!}{\gamma! \delta!} x^{\gamma} (1-x)^{\delta} \quad , \quad \gamma > -1, \delta > -1, 0 \leq x \leq 1 \quad (2.53)$$

$$\mu_r = \frac{(\gamma + \delta + 1)!(\gamma + r)!}{(\gamma + \delta + r + 1)! \gamma!} \quad , \quad r=1,2,\dots,N \quad (2.54)$$

もちろん，式(2.51)，(2.53)は，それぞれがもつパラメタの値により多くの形状が表現され得る。

3.1.2 適用性の検討

ここでは，本章2.4の方法に従って，1変数最大エントロピー分布の適用性の検討を行う。

いま， $G(3,1)$ の場合に，式(2.30)により，式(2.52)を制約条件として式(2.51)を適合させた結果が図2.4であり，同様に， $B(2,2)$ の場合に，式(2.54)を制約条件として式(2.53)を適合させた結果が図2.5である。ここで， $G(3,1)$ の3モーメント適合と5モーメント適合は，式(2.36)の存在条件を満たさなかったため，得ることができなかった。

次に， $G(3,1)$ より左上にピークを持つ $G(2,1)$ ，さらに左上にピークが移った $G(1,1)$ ，および $B(2,2)$ よりもピーク値の低い $B(1,1)$ についても同様に検討した結果，4次モーメントまでを制約条件とした1変数最大エントロピー分布で，ほぼ母集団の形状が記述されることがわかった。さらに，水文統計上重要な関心事である母集団の右裾・左裾も，4モーメント適合になるとよく再現でき，実用的に十分な精度が得られていると判断できた。

また，表2.1は，図2.4，2.5で示した1変数最大エントロピー分布のパラメタの値である。

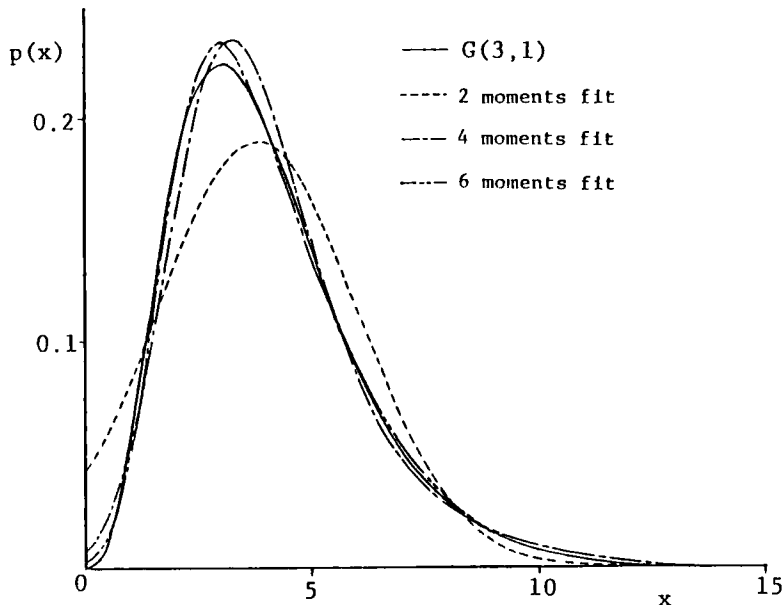


図2.4 $G(3,1)$ の最大エントロピー適合

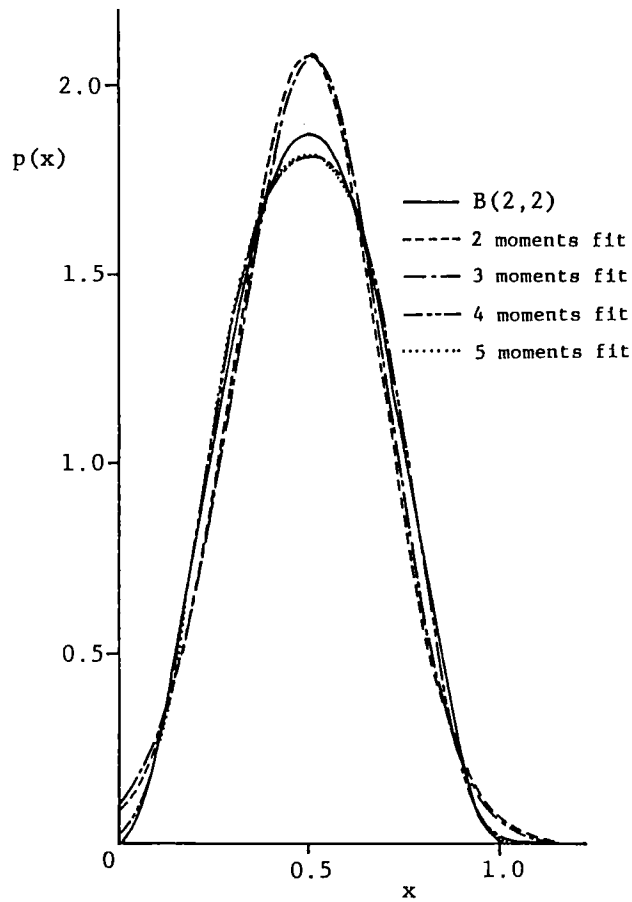


図2.5 B(2,2)の最大エントロピー適合

表2.1 G(3,1)とB(2,2)に対する最大エントロピー分布のパラメタの値

Dist.	Number of moments	λ_0	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
G(3,1)	2	1.77	-3.15	1.66				
	4	3.43	-10.1	9.72	-3.31	0.413		
	6	4.54	-17.3	24.9	-17.2	6.51	-1.24	0.0929
B(2,2)	2	3.29	-6.70	3.37				
	3	3.03	-5.63	2.19	0.382			
	4	4.67	-16.0	21.8	-13.8	3.46		
	5	4.51	-14.6	17.8	-9.07	0.924	0.502	

x の最高次数にかかる λ の値がすべて正値をとっている。このことから、ここでも求まった1変数最大エントロピー分布が、式(2.28)を満たすための必要条件を満足していることがわかる。

なお、ここでは水文量の頻度分析上重要な分布の1つであるGamma分布と、正の領域で対称形を示すBeta分布に対する適合度をみるため、上述の計算結果を示したが、文献4)、6)にも他の非対称確率分布に対する同種の例があることを付記する。

3.2 モーメントの安定性と感度分析

3.2.1 モーメントの安定性

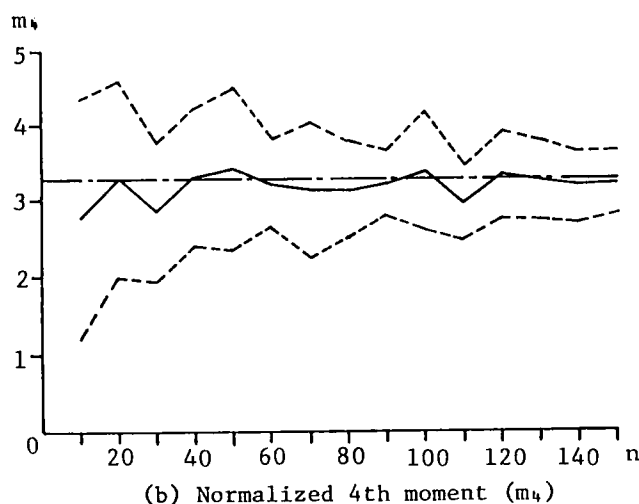
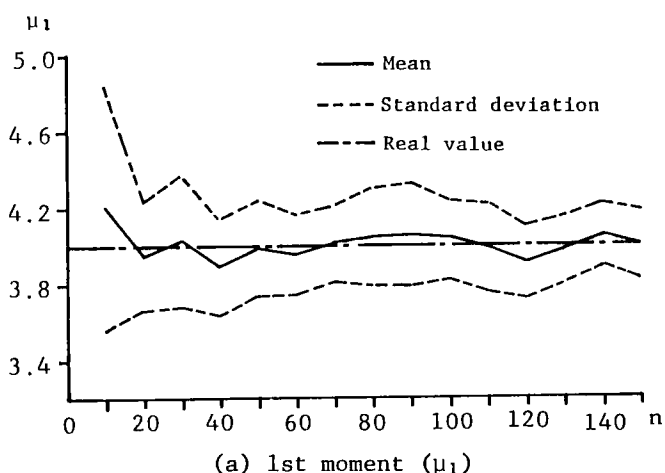


図2.6 データ数の増加に伴う $G(3,1)$ のモーメントの安定性

さて、一般に実現値から得られるモーメントは、母集団のモーメントに一致せず、モーメントの次数が大きくなるに従って、真値からのずれが大きくなることが予想される。当然、それは推定分布にひずみを生じさせることになる。

そこで、まず $G(1,1)$, $G(2,1)$, $G(3,1)$, $B(1,1)$ および $B(2,2)$ を母集団として、個数 n ($n=10, 20, \dots, 150$) のデータを20回発生させて、各個数の6次までのモーメントの平均値と標準偏差を求めることにより、モーメントの安定性を評価した。図2.6, 2.7は、1例として3.1 で採用した非対称分布の $G(3,1)$ と対称分布の $B(2,2)$ について、確率変数の平均値の大きさと変動をみ

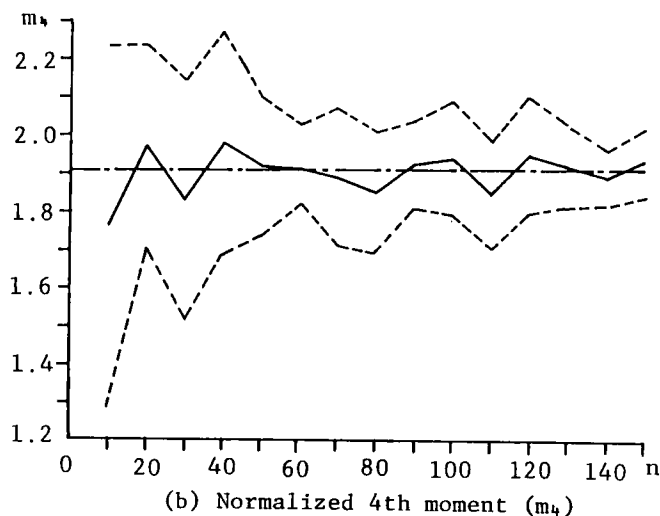
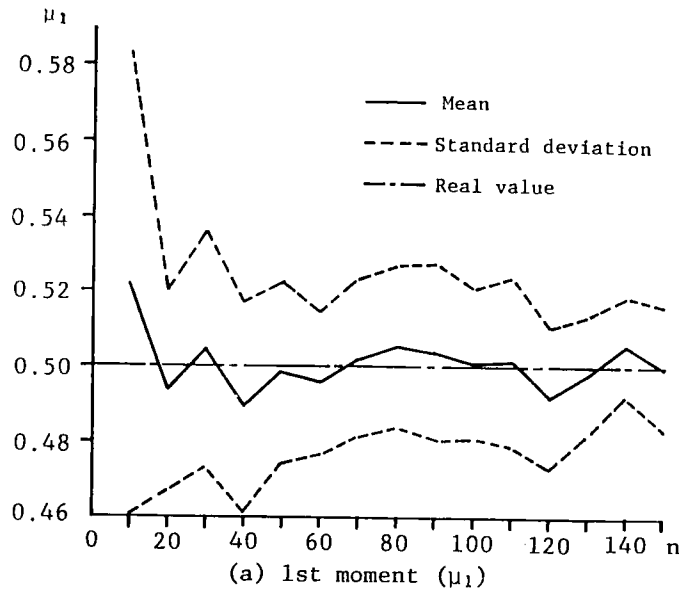


図2.7 データ数の増加に伴う $B(2,2)$ のモーメントの安定性

るための1次モーメントの安定性を、およびとがり度に関係する基準化された4次モーメントの安定性を示したものである。これらの図より、モーメントの次数が小さい程、真値に対する相対的な変動が小さく、データ数の増加とともに、平均値が母集団のモーメントに近づき、標準偏差が小さくなっていく様子がうかがわれる。また、他の母集団の場合も同様の傾向を示した。

3.2.2 モーメントの変動による感度分析

3.1.2 より、4モーメント適合が母集団の形状をよく記述できることがわかったため、ここでは、その最高次数モーメント m_4 の変動が1変数最大エントロピー分布に及ぼす影響を分析する。

いま、 $G(3,1)$ を母集団として、3次モーメントまで真のモーメントを採用し、 m_4 のみを推定分布が存在する区間 ($3.25 \leq m_4 \leq 4.50$ 程度) で変化させて、最大エントロピー分布のパラメタ同定を行った。

表2.2 は、その時のパラメタの値であり、 m_4 の増加に従って、 $\lambda_0, \lambda_2, \lambda_4$ は減少し、 λ_1, λ_3 は増加していることがわかる。この表の結果を用いて、得られた密度関数を図化したのが図2.8 である。 $3.25 \leq m_4 \leq 4.50$ の区間では、得られた分布は図2.4 の2モーメント適合より母集団に対してよい適合性を示しており、 m_4 の変動に伴う最大エントロピー分布の感度は、比較的鈍いことがわかる。また、右裾の部分は上述の区間でよく母集団を再現しているが、左裾の部分ではその区間が $3.25 \leq m_4 \leq 3.30$ 程度にせばまっている。一方、図2.6(b)を参考にすると、母集団を $G(3,1)$ とした場合、4次モーメントまで採用するには、データ数が60個程度必要であることが推察される。ただし、ここでの考察は、3次モーメントまでは真値に固定しているが、実データから得られるモーメントには当然それらにも誤差が含まれていることに注意を要する。

表2.2 $G(3,1)$ における m_4 の変動と最大エントロピー分布
のパラメタの値との関係

$m_4 \backslash \lambda_i$	λ_0	λ_1	λ_2	λ_3	λ_4
3.25	4.24	-13.9	15.1	-6.13	0.901
•3.28	3.43	-10.1	9.69	-3.30	0.410
3.30	3.14	-8.73	7.89	-2.40	0.264
3.50	2.41	-5.61	4.13	-0.771	0.0450
4.00	2.09	-4.39	2.89	-0.361	0.0130
4.50	1.98	-3.96	2.46	-0.228	0.0058

• : Real value

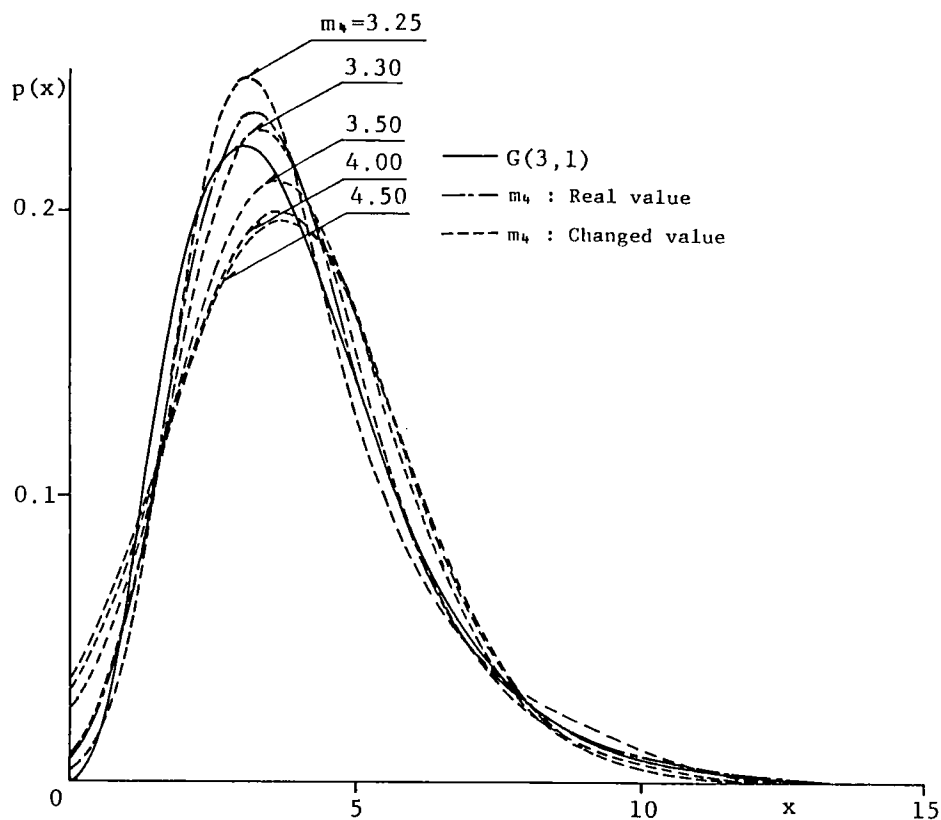


図2.8 $G(3,1)$ における m_4 の変動と最大エントロピー分布

第4節 実データへの適用

4.1 実データの種類の抽出法

水文量として降水量と流量を採用した。一般に、これらの水文量は、時間単位の取り方により、逆J字型、指数型、対数正規型、正規型等の分布形を示すことが知られている。^{14), 15)}そこで、ここでは多くの形状の頻度図への1変数最大エントロピー分布の適用性をみるため、種々の時間単位の水文量を取り上げた。

一方、データ数とモーメントの変動関係を調べると、明確ではないが1次、2次モーメントではデータ数が40個付近から、3次、4次モーメントでは60個付近から安定する傾向がみられる。しかし、採用した観測所の観測期間はこれらの値より短い場合が多く、データ不足が否めないため、データ間の独立性をある程度犠牲にしてデータ数の増加をはかった箇所もある。

4.1.1 降水データ

長野県下の古谷（採用期間25年），北牧（29年），長久保新町（26年），福島（42年）および藪原（42年）の降水データをデータ数を考慮して次のような時間単位で採用した。ただし，年最大日降水量については，上に掲げた観測所では十分なデータ数が得られず，また上述のようにしてデータ数の増加をはかることができないため，大阪のデータ（72年）を用いた。

- 1) 3時間目：主降雨時間帯開始時点から3時間目の時間降水量($t=3$)。
- 2) 6時間目：主降雨時間帯開始時点から6時間目の時間降水量($t=6$)。
- 3) 6時間合計：主降雨の1時間目から6時間目までの合計。
- 4) 12時間合計：主降雨の1時間目から12時間目までの合計。

ここに，主降雨時間帯とは，出水をもたらす一連降水（今回は総降水量30mm以上）によって，流量が流域末端で立ち上がる時刻から15時間分とする。ただし，福島，藪原は時間降水量のデータが得られないため，上記時間単位の検討は行わなかった。

5) 日：古谷，北牧および長久保新町は6月11日から6月20日までの10日間で3mm以上の場合を日（6月），9月11日から9月20日までの10日間で3mm以上の場合を日（9月）とする。福島，藪原は6月15日を日（6月），9月15日を日（9月）とし，3mm以下の場合にはその前後の3mm以上の降水日を採用する。

6) 旬：古谷，北牧および長久保新町は，6月上旬から7月下旬までの6旬を旬（6，7月），9月上旬から10月下旬までの6旬を旬（9，10月）とし，福島，藪原は6月中旬を旬（6月），9月中旬を旬（9月）とする。

7) 月：古谷，北牧および長久保新町は，5，6，7月の3ヶ月を月（5，6，7月），8，9，10月の3ヶ月を月（8，9，10月）とし，福島，藪原は6月を月（6月），9月を月（9月）とする。

8) 季節：どの観測所も，3，4，5月の合計を季節（春），6，7，8月の合計を季節（夏），9，10，11月の合計を季節（秋），12，1，2月の合計を季節（冬）とする。

9) 年：年総降水量。

10) 年最大日：年間を通じての最大日降水量。

11) 年最小月：年間を通じての最小月降水量。

4.1.2 流量データ

長野県下の生田（採用期間6年），杭瀬下（28年），立ヶ花（30年）および小市（27年）の流量データをデータ数を考慮して次のような時間単位で採用した。ただし，生田では観測期間が短いため日以上での適用は行わず，小市では日単位以上のデータしか得ることができなかった。また，単位は $[m^3/s]$ に統一した。

- 1) 3時間目：主流量時間帯開始時点から3時間目の時間流量($t=3$)。
- 2) 6時間目：主流量時間帯開始時点から6時間目の時間流量($t=6$)。
- 3) 6時間平均：主流量の1時間目から6時間目までの平均時間流量。
- 4) 12時間平均：主流量の1時間目から12時間目までの平均時間流量。

ここに、主流量時間帯とは、主降雨によるハイドログラフの立ち上がる時刻から15時間分とする。

5) 日：6月15日を日(6月15日)，9月15日を日(9月15日)とする。また，データ数を増加させるため，6月11日から6月20日までの10日間を日(6月)，9月11日から9月20日までの10日間を日(9月)として採用する。

6) 旬：6月上旬から7月下旬までの6旬を旬(6，7月)，9月上旬から10月下旬までの6旬を旬(9，10月)とする

7) 月：5，6，7月の3ヶ月を月(5，6，7月)，8，9，10月の3ヶ月を月(8，9，10月)とする。

8) 年：年総流量。

9) 年最大日：年間を通じての最大日流量。

10) 年最大月：年間を通じての最大月流量。

11) 年最小日：年間を通じての最小日流量。

12) 年最小月：年間を通じての最小月流量。

4.2 適用結果と考察

本項では，4.1で述べた実データを用い，2.4の方法に従って，平均，分散を情報として取り入れた2モーメント適合と，さらに，ひずみ度，とがり度を情報として加えた4モーメント適合を得た。以下，その結果について考察する。

4.2.1 パラメタの値

表2.3は，基準化されたモーメントから算定された推定分布のパラメタである(表中の空白部分はパラメタが同定できなかったことを示す)。これより，次のことが読み取れる。

1) 2モーメント適合は，理論的には $1 \leq m_2 \leq 2$ となる場合に存在するが，ここで用いた計算法では，存在区間内でも長久保新町(時間単位：年)の $m_2=1.0210$ ，および立ヶ花(年)の $m_2=1.0208$ の場合に得ることができなかった。

2) 2モーメント適合のパラメタの値は，正負交互に現われることが多く， $\lambda_2 > 0$ の収束条件(式(2.28)を満たすための必要条件)を満足している。 $m_2 > 1.6$ の区間では，すべてのパラメ

表2.3 ヒストグラムに対する最大エントロピー分布のパラメタの値

(a) Koya (Precipitation)

Unit time	μ_1 (mm)	λ_0	λ_1	λ_2	λ_3	λ_4
One hour (t=3)	2.7353	0.7090 1.3814	-0.6072 -4.2364	0.5497 5.3438	-2.1986	0.3208
One hour (t=6)	2.6304	0.3449 0.7371	0.2467 -2.0822	0.2308 3.3247	-1.3692	0.1887
Six hours	16.309	0.9366 1.4011	-1.1513 -4.4682	0.7716 6.4486	-3.3572	0.6305
Twelve hours	28.063	0.2103 1.9417	0.5512 -6.5578	0.1296 7.0487	-2.2069	0.2186
One day (Jun.)	14.429	0.3497 1.5672	0.2358 -4.7424	0.2346 5.0132	-1.5051	0.1455
One day (Sep.)	20.426	-0.1377 0.2670	0.8741 0.4240	0.3263 0.1708	-0.0667	0.0031
Ten days (Jun., Jul.)	48.749	0.7274	-1.2374	1.4347	-0.2890	0.0181
Ten days (Sep., Oct.)	44.974	0.0086 0.1721	0.9827 0.3319	0.0044 0.5125	-0.1215	0.0086
One month (May, Jun., Jul.)	129.14	1.7607 3.9587	-3.1272 -12.4192	1.6496 12.8344	-4.8617	0.6649
One month (Aug., Sep., Oct.)	134.80	0.7236 2.2082	-0.6420 -6.2655	0.5635 5.8417	-1.6236	0.1471
Season (Spring)	236.06	8.5578	-18.1781	9.0892		
Season (Summer)	436.56	4.8627	-10.2219	5.1148		
Season (Fall)	304.57	2.7269 11.5134	-5.3999 -36.5661	2.7327 37.7375	-14.9212	2.0548
Season (Winter)	88.662	1.8593 6.0600	-3.3618 -22.2848	1.7588 27.3558	-13.0776	2.2166
Year	1069.9	10.0279	-21.2756	10.6378		
Annual minimum one month	10.225	0.5475 1.3177	-0.2247 -4.5836	0.4015 6.3235	-2.7648	0.4094

表2.3 (続き)

(b) Kitamaki (Precipitation)

Unit time	μ_1 (mm)	λ_0	λ_1	λ_2	λ_3	λ_4
One hour (t=3)	2.8125	0.6350	-0.4313	0.4806		
One hour (t=6)	2.3333	0.8432	-1.9410	2.9156	-1.3122	0.2201
Six hours	16.233	0.3325	0.2751	0.2210		
		1.7043	-6.3795	7.9886	-3.0680	0.3821
		2.0073	-3.7130	1.9237		
Twelve hours	28.481	5.6980	-24.2921	36.2399	-21.7729	4.6316
		2.2843	-4.3666	2.2342		
		5.7157	-20.5618	25.5007	-12.7880	2.3531
One day (Jun.)	13.957	0.0009	0.9982	0.0005		
		1.0433	-2.7642	2.8700	-0.6621	0.0459
One day (Sep.)	17.453	0.3119	-0.4679	1.1955	-0.2598	0.0159
Ten days (Jun., Jul.)	47.087	0.1859	0.6053	0.1125		
		0.5137	-0.6005	1.0228	-0.2071	0.0131
Ten days (Sep., Oct.)	44.825	0.2103	0.5512	0.1296		
		0.2455	0.0942	0.9469	-0.4140	0.0603
One month (May, Jun., Jul.)	127.77	1.4853	-2.4689	1.3477		
		2.5397	-7.8562	8.8465	-3.7914	0.6222
One month (Aug., Sep., Oct.)	118.49	1.2376	-1.8743	1.0822		
		2.4757	-8.8317	11.5453	-5.6935	1.0094
Season (Spring)	242.17	5.5435	-11.7148	5.8593		
Season (Summer)	383.29	6.2769	-13.3064	6.6541		
		8.4008	-20.6966	14.7221	-3.0049	0.1914
Season (Fall)	300.93	3.9989	-8.3009	4.1596		
		5.3050	-14.7329	14.2842	-6.3103	1.3478
Season (Winter)	95.404	1.6149	-2.7791	1.4890		
		3.7008	-12.5547	14.4925	-6.3531	1.0101
Year	1026.3	13.9341	-29.4117	14.7059		
Annual minimum one month	14.114	1.4322	-2.3416	1.2902		

表2.3 (続き)

(c) Nagakuboshinmachi (Precipitation)

Unit time	μ_1 (mm)	λ_0	λ_1	λ_2	λ_3	λ_4
One hour (t=3)	3.2317	0.2540 1.4670	0.4534 -4.3254	0.1612 4.4833	-1.2599	0.1120
One hour (t=6)	2.2167	0.4629 1.7041	-0.0263 -6.7823	0.3279 9.2975	-4.0894	0.5922
Six hours	17.271	0.4728 2.0510	-0.0495 -6.4762	0.3364 6.7268	-2.1178	0.2167
Twelve hours	29.690	2.0497 5.0432	-3.8135 -18.8138	1.9711 24.4105	-12.7477	2.4202
One day (Jun.)	17.087	0.2515 1.0205	0.4591 -3.1234	0.1593 3.9488	-1.3252	0.1439
One day (Sep.)	17.485	0.0209 0.8717	0.9578 -2.0238	0.0108 2.1752	-0.4704	0.0304
Ten days (Jun., Jul.)	54.654	0.3791 0.6424	0.1681 -0.8165	0.2582 1.0354	-0.1813	0.0103
Ten days (Sep., Oct.)	46.179	0.1946 0.3387	0.5861 -0.2159	0.1185 1.0580	-0.3560	0.0411
One month (May, Jun., Jul.)	143.09	2.2097 2.6924	-4.1911 -6.1962	2.1504 4.4014	-0.8254	0.0676
One month (Aug., Sep., Oct.)	125.59	1.3632 3.3018	-2.1761 -12.4793	1.2160 16.4159	-8.2211	1.4586
Season (Spring)	260.21	5.8871	-12.4625	6.2327		
Season (Summer)	445.36	8.4042	-17.8529	8.9266		
Season (Fall)	304.69	5.7446 6.0744	-12.1529 -13.3659	6.0781 7.4167	-0.4609	0.0062
Season (Winter)	106.57	3.3363 7.1313	-4.4887 -25.1234	2.2927 29.7169	-13.9439	2.3593
Year	1116.8					
Annual minimum one month	15.587	0.9175 2.1881	-1.1054 -7.5053	0.7525 9.1138	-3.8667	0.5741

表2.3 (続き)

(d) Fukushima (Precipitation)

Unit time	μ_1 (mm)	λ_0	λ_1	λ_2	λ_3	λ_4
One day (Jun.)	26.483	-0.3689	1.7438	0.0019	-0.1118	0.0148
One day (Sep.)	17.376	0.0840	0.8265	0.0465		
Ten days (Jun.)	86.812	1.0108	-3.4215	4.3594	-1.4154	0.1434
Ten days (Sep.)	73.424	0.2304	0.5065	0.1439		
One month (Jun.)	266.57	1.0016	-2.9998	3.7418	-1.2141	0.1269
One month (Sep.)	229.88	0.6615	-0.4942	0.5051		
Season (Spring)	472.04	1.1745	-4.0268	6.1838	-3.1071	0.5360
Season (Summer)	702.23	2.6437	-5.2068	2.6389		
Season (Fall)	465.61	6.5292	-20.4223	20.8507	-8.1227	1.1472
Season (Winter)	194.35	1.7268	-3.0462	1.6120		
Year	1821.9	2.0928	-5.0106	4.4618	-1.5029	0.2570
Annual minimum one month	30.702	4.7177	-9.9017	4.9553		
		21.7962	-66.8704	68.3605	-27.7130	3.9628
		7.4683	-15.8634	7.9320		
		6.1149	-12.9561	6.4792		
		2.4462	-4.7461	2.4163		
		6.6393	-25.2595	33.4780	-18.2014	3.6041
		16.3901	-34.4828	17.2414		
		0.7820	-0.7810	0.6193		
		1.3470	-4.1203	5.4442	-2.4281	0.3898

表2.3 (続き)

(e) Yabuhara (Precipitation)

Unit time	μ_1 (mm)	λ_0	λ_1	λ_2	λ_3	λ_4
One day (Jun.)	18.979	0.0965	-0.0117	1.0207	-0.2283	0.0136
One day (Sep.)	20.269	0.2451	0.4734	0.1546		
		1.1965	-3.8461	4.6550	-1.5497	0.1659
Ten days (Jun.)	89.344	0.2657	0.4270	0.1698		
		0.5616	-0.9142	1.4925	-0.4246	0.0415
Ten days (Sep.)	91.026	0.6324	-0.4252	0.4783		
		0.9689	-1.9349	2.0511	-0.5403	0.0552
One month (Jun.)	279.56	2.1739	-4.1066	2.1102		
		7.4417	-27.5365	34.3097	-16.9937	3.0060
One month (Sep.)	237.41	2.1161	-3.9702	2.0453		
		3.2393	-8.9645	8.3407	-2.8604	0.4030
Season (Spring)	437.96	4.2137	-8.7817	4.3983		
Season (Summer)	731.63	4.3296	-9.0403	4.5267		
		8.7572	-23.3635	19.1335	-5.3645	0.5133
Season (Fall)	478.82	7.2386	-15.3729	7.6868		
Season (Winter)	226.01	2.4180	-4.6800	2.3845		
Year	1887.0	11.2385	-23.8092	11.9046		
Annual minimum one month	33.338	0.4891	-0.0875	0.3504		
		0.5950	-1.0214	1.9676	-0.8878	0.1475

表2.3 (続き)

(f) Osaka (Annual maximum daily rainfall)

μ_1 (mm)	λ_0	λ_1	λ_2	λ_3	λ_4
95.774	3.3376	-6.8047	3.4201		
	11.5330	-35.7668	36.1557	-14.0823	1.9323

表2.3 (続き)

(g) Ikuta (Discharge)

Unit time	μ_1 (m ³ /s)	λ_0	λ_1	λ_2	λ_3	λ_4
One hour (t=3)	91.462	1.3816 5.3392	-2.2202 -16.7213	1.2357 16.3439		
One hour (t=6)	143.32	0.8179 3.2833	-0.8669 -11.8189	0.6541 13.6443	-5.5587 -5.4391	0.6337 0.7288
Six hours	101.73	1.2703 4.3251	-1.9530 -14.0038	1.1169 14.3828		
Twelve hours	133.39	1.0220 3.3314	-1.3563 -11.2596	0.8581 12.3214	-5.1727 -4.6869	0.6320 0.6079

表2.3 (続き)

(h) Kuissege (Discharge)

Unit time	μ_1 (m ³ /s)	λ_0	λ_1	λ_2	λ_3	λ_4
One hour (t=3)	102.79	2.3630 9.4826	-4.5513 -30.2015	2.3227 30.9764	-11.9780	1.6040
One hour (t=6)	146.44	1.7694 4.9807	-3.1478 -15.4048	1.6591 15.1418	-5.2956	0.6389
Six hours	109.73	2.3496 7.4520	-4.5199 -22.7594	2.3076 22.0614	-7.8105	0.9517
Twelve hours	151.36	1.4718 3.5445	-2.4365 -10.3555	1.3330 9.6039	-2.9653	0.3113
One day (Jun.15)	88.67	0.1793	-0.6125	1.7093	-0.4232	0.0288
One day (Jun.)	75.35	0.3825	-0.4696	1.0036	-0.1778	0.0087
One day (Sep.15)	63.03	0.3434 2.8904	0.2503 -9.5092	0.2296 9.5451	-2.9395	0.2882
One day (Sep.)	83.79	0.1225	0.4112	0.3762	-0.0474	0.0015
Ten days (Jun.,Jul.)	77.97	0.6807 2.2217	-0.5398 -6.7717	0.5230 6.8226	-2.1378	0.2217
Ten days (Sep.,Oct.)	77.85	0.5230 2.1165	-0.1671 -6.2471	0.3799 6.0269	-1.7166	0.1575
One month (May,Jun.,Jul.)	70.67	2.1274 5.8644	-3.9970 -22.1816	2.0581 28.9042	-15.1579	2.8710
One month (Aug.,Sep.,Oct.)	69.61	1.0436 3.4104	-1.4082 -11.7709	0.8802 13.2025	-5.2104	0.7036
Year	57.38	8.8828	-18.8652	9.4327		
Annual maximum one day	617.35	0.7207 4.8552	-0.6350 -20.5948	0.5607 27.2942	-12.8845	2.0212
Annual maximum one month	121.49	5.1808 8.3025	-10.9216 -21.9076	5.4636 17.6368	-4.7966	0.4520
Annual minimum one day	16.44	2.4040	-4.6474	2.3689		
Annual minimum one month	26.65	7.5888	-16.1204	8.0604		

表2.3 (続き)

(i) Tategahana (Discharge)

Unit time	μ_1 (m ³ /s)	λ_0	λ_1	λ_2	λ_3	λ_4
One hour (t=3)	416.47	2.2716 6.0428	-4.3367 -19.2719	2.2199 20.0108	-7.8442	1.0994
One hour (t=6)	509.07	1.9165 5.4364	-3.4976 -18.6382	1.8224 21.1599	-9.2111	1.4326
Six hours	427.97	2.2843 6.0539	-4.3666 -19.3840	2.2342 20.2523	-8.0178	1.1375
Twelve hours	502.12	2.1048 5.3702	-3.9435 -18.4658	2.0327 21.3763	-9.6993	1.6014
One day (Jun.15)	292.80	0.2451 4.7554	0.4734 -17.4528	0.1546 18.8907	-6.5960	0.7340
One day (Jun.)	281.66	0.2566 1.9326	0.4476 -4.9578	0.1631 4.0967	-0.8429	0.0524
One day (Sep.15)	262.08	1.2985 10.3590	-2.0206 -36.6045	1.1467 41.2519	-17.3492	2.4714
One day (Sep.)	292.93	0.2527 2.0297	0.4563 -5.3626	0.1602 4.5049	-0.9688	0.0634
Ten days (Jun.,Jul.)	330.69	1.4066 6.5988	-2.2802 -24.1617	1.2626 28.5303	-12.6185	1.9173
Ten days (Sep.,Oct.)	267.72	1.4718 5.3824	-2.4365 -17.4905	1.3330 18.0027	-6.6338	0.8329
One month (May,Jun.,Jul.)	299.29	3.5162 12.2590	-7.2112 -41.8790	3.6205 48.6833	-23.1788	4.0367
One month (Aug.,Sep.,Oct.)	248.17	2.6116	-5.1321	2.6028		
Year	229.47					
Annual maximum one day	1550.6	2.7440 11.1695	-5.4397 -40.8823	2.7520 50.7916	-25.5952	4.6312
Annual maximum one month	432.89	15.8323	-33.3333	16.6667		
Annual minimum one day	101.71	12.4417	-26.3157	13.1578		
Annual minimum one month	124.56	16.3901	-34.4828	17.2414		

表2.3 (続き)

(j) Koichi (Discharge)

Unit time	μ_1 (m ³ /s)	λ_0	λ_1	λ_2	λ_3	λ_4
One day (Jun.15)	171.65	0.2553 2.2066	0.4505 -8.2015	0.1621 9.5071	-3.3804	0.3842
One day (Jun.)	163.65	0.1191 1.3061	0.7513 -3.1100	0.0682 2.7464	-0.5401	0.0317
One day (Sep.15)	125.62	1.4718 8.2457	-2.4365 -26.3585	1.3330 26.4132	-9.5282	1.1529
One day (Sep.)	151.05	0.2103 1.4564	0.5512 -3.5185	0.1296 3.0308	-0.6074	0.0367
Ten days (Jun.,Jul.)	191.39	1.0839 4.9210	-1.5050 -18.0441	0.9216 21.0937	-8.9227	1.2805
Ten days (Sep.,Oct.)	135.72	1.9165 7.1662	-3.4976 -23.2182	1.8224 24.0639	-9.2338	1.2220
One month (May,Jun.,Jul.)	175.15	2.8695 10.4144	-5.7302 -36.2843	2.8935 43.1730	-21.0825	3.7781
One month (Aug.,Sep.,Oct.)	128.12	3.2427 12.5171	-6.5881 -46.1385	3.3136 57.9593	-29.6939	5.4647
Year	130.29	8.5478	-18.1781	9.0892		
Annual maximum one day	886.58	3.3621	-6.8606	3.4476		
Annual maximum one month	248.45	7.2386	-15.3728	7.6868		
Annual minimum one day	47.524	10.4792	-22.2216	11.1108		
Annual minimum one month	64.501	18.3749	-38.4615	19.2307		

タは正值となり、この場合2モーメント適合のピークは、 $x = 0$ 上に存在して指数分布に近い形となる。

3) 4モーメント適合は、 m_2 、 m_3 、 m_4 に従属するため、その存在区間を明示することが難しいが、 $m_4 < 2$ のとき存在しない場合が多い。

4) 4モーメント適合のパラメタも、正負交互に現われることが多く、 $\lambda_4 > 0$ の収束条件を満足している。正負交互の並びが乱れる場合は、 $m_4 > 10$ となっているが、この m_4 の条件下で上述の並びを保持している場合も多数存在する。

4.2.2 ヒストグラムと推定分布

図2.9に、ヒストグラムと表2.3のパラメタを用いて得られた推定分布を示す。ここでは紙面の都合上、古谷（降水量）と大阪（降水量）のみに留めるが、他の分布も同様な傾向を持つ。この図より以下のことが考察される。

1) 降雨データの場合、時間単位毎のヒストグラムをみると、時間、日、旬単位ではピークが左寄りになり右に長く尾引く形状になるが、月、季節、年になるに従いヒストグラムのピークが右に移り左右対称に近づく。これは従来知見に一致する。流量データでは、上述のような顕著な傾向はみられない。ヒストグラムの中には、明確なピークがみられない場合があり、連続性のとぎれることが多い。

2) 全体的には、2モーメント適合から4モーメント適合になると適合度は大きく改善される。すなわち、ひずみ度、とがり度の情報が有効に作用していることがわかる。

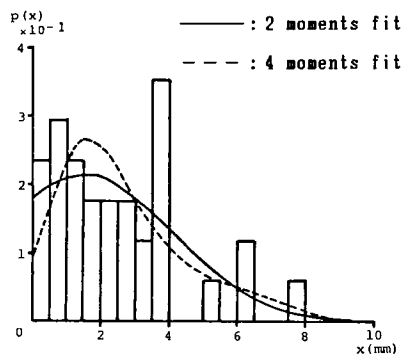
3) 2モーメント適合はなめらかな形状を示し、ヒストグラムより偏平になる場合が多い。

4) 4モーメント適合は、ヒストグラムのピークに分布形のピークが近づき、ヒストグラムのピークが2つ現われる場合（第7章 2.3で詳述）も説明できる。しかし、この適合は原点から離れたヒストグラムに対して得られない場合が多い。

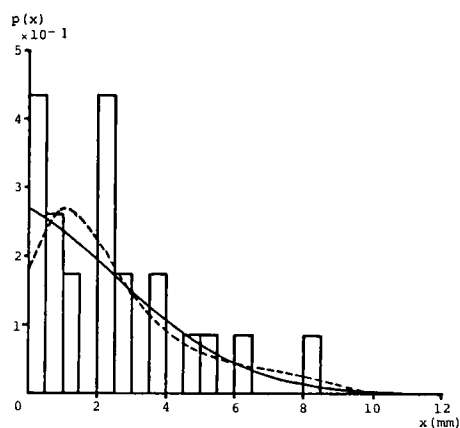
5) 4モーメント適合にすると、左裾の適合度の改善が明確に現れる場合が多く、また右裾に離れて存在するデータのヒストグラムも説明できるように、分布形状が長く尾を引くように移行していることがわかる。

6) 4モーメント適合が得られないとき、2モーメント適合の分布形は、ヒストグラムの傾向を比較的良好に説明し、ピークが明確で正規分布に近い形をとっている。一方、 $m_2 > 2$ で2モーメント適合が存在しないとき、4モーメント適合の分布形は指数分布に近い形となっているが、 $m_2 \approx 1$ で2モーメント適合が得られないとき、4モーメント適合も得られなかった。

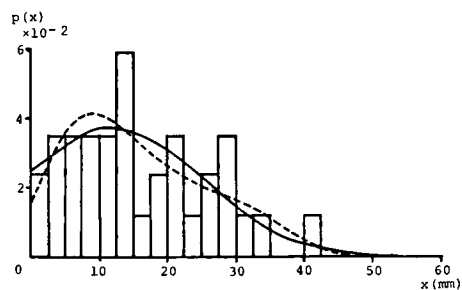
ところが、一方現実には、ここで取り上げたデータ数より極めて小標本で、例えば、10個、20個程度の実データで統計解析を要求される場合がある。このようなとき、高次モーメントにおける推定誤差の増大が予想されるため、まず第1に近隣の観測所の長期間の水文データから不



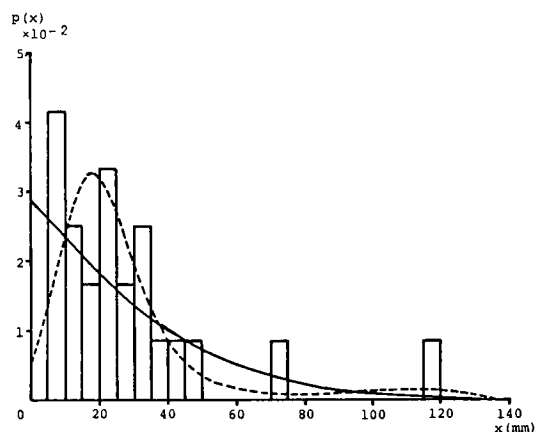
(a) One hour ($t=3$)



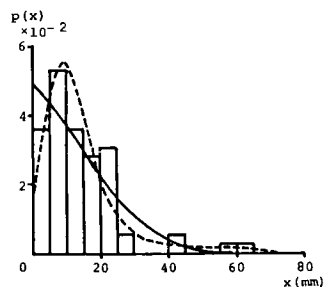
(b) One hour ($t=6$)



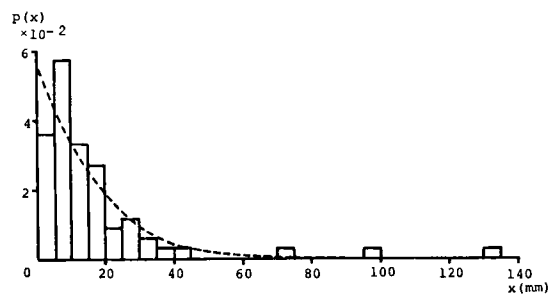
(c) Six hours



(d) Twelve hours

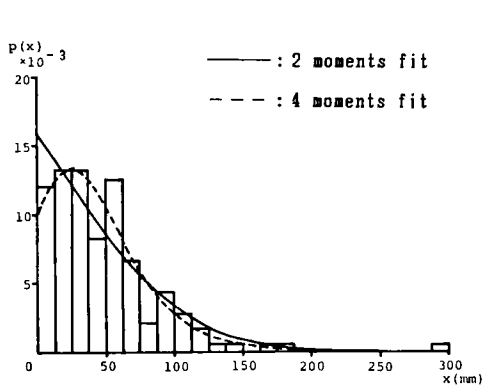


(e) One day (Jun.)

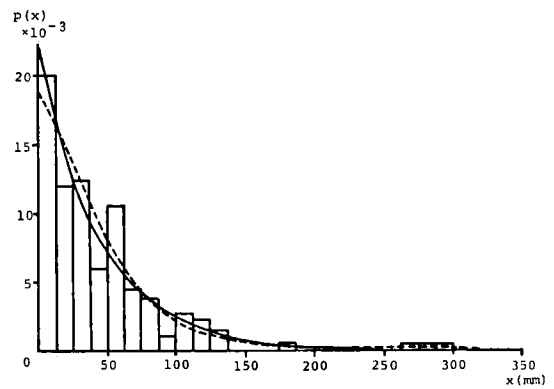


(f) One day (Sep.)

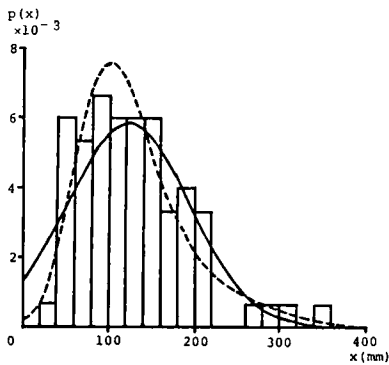
図2.9 ヒストグラムと最大エントロピー分布



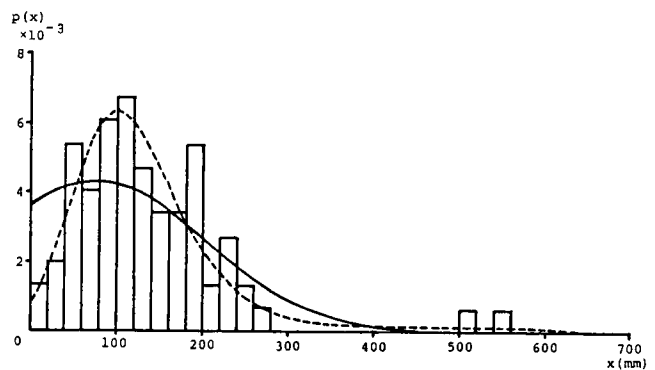
(g) Ten days (Jun., Jul.)



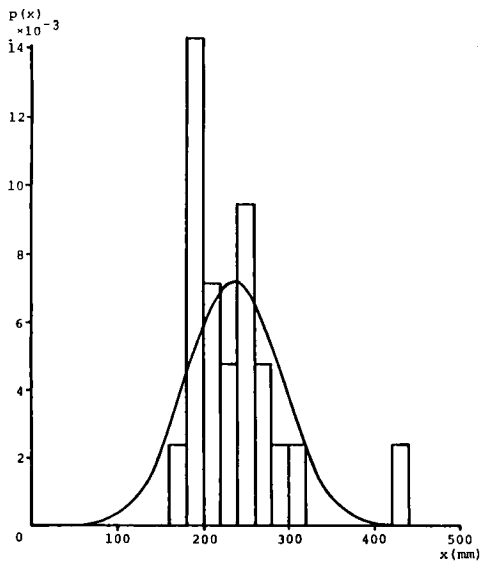
(h) Ten days (Sep., Oct.)



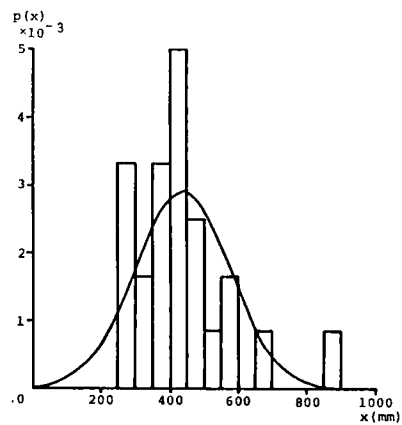
(i) One month (May, Jun., Jul.)



(j) One month (Aug., Sep., Oct.)

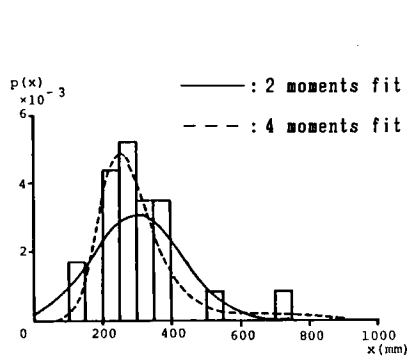


(k) Season (Spring)

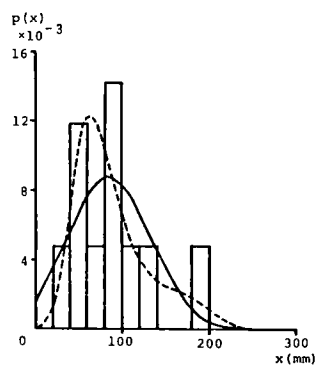


(l) Season (Summer)

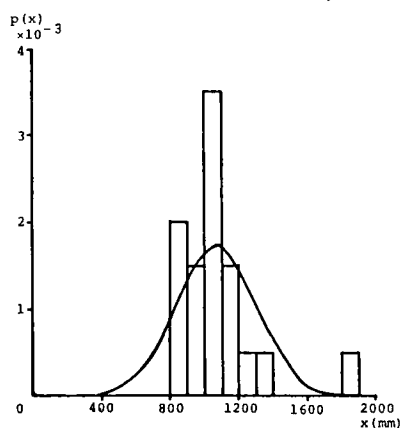
図2.9 (続き)



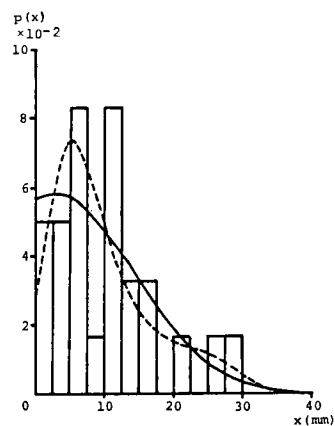
(m) Season (Fall)



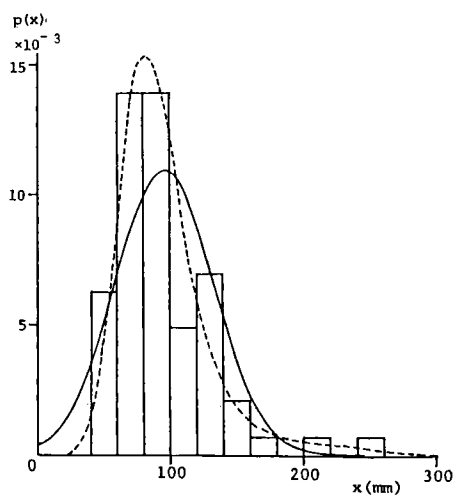
(n) Season (Winter)



(o) Year



(p) Annual minimum
one month



(q) Annual maximum
one day

図2.9 (続き)

足する観測所のデータを補充することを考えるべきであろう。しかし、それが精度的に不可能ならば、3.2.1の結果からみると2次モーメントまでは、3次、4次モーメントと比較して相対的に安定しているので、得られた結果の信頼性は十分でないが、統計モーメントだけを情報とするなら2モーメント適合を用いればよかろう。ただし、著者は現時点では、このようなデータ数では、第6章で述べるように統計モーメントと指数関数の期待値の両方を情報とした分布を用いる方が、2モーメント適合の場合より精度の高い確率水文量あるいはリターン・ピリオドを得ることができると考えている。

第5節 要 約

確率分布を用いた統計解析では、1変数分布が基礎となるため、本章では1変数最大エントロピー分布について詳細に論じた。得られた成果を要約すると次のようになる。

1) 1変数最大エントロピー分布の定式化について説明しながら、最大エントロピー分布による水文量の頻度分析の妥当性を、この分布のもつ特性と水文データの小標本性から明らかにした。

2) 1変数最大エントロピー分布から統計学においてよく知られている既存分布が導出できることを紹介し、これを用いて、この分布がPearson系の分布等と同レベルで位置付けされる分布系であることを明確にした。

3) 1変数最大エントロピー分布の、パラメタ同定法、解の存在条件、モーメントの基準化は、具体的に統計モーメントを与えて最大エントロピー分布を同定するときに重要となるので、要約して説明しておいた。

4) いくつかの、典型的な形状を示す1変数Gamma分布と1変数Beta分布を母集団と仮定し、真のモーメントを情報として用いた1変数最大エントロピー分布で適合度を検討して、4次モーメントまで情報とすると十分よい再現性が得られることを示した。

5) ついで、モーメントの安定性について、シミュレーションデータを用いて調べ、この結果と4)の検討を踏まえて、4次モーメントの推定誤差が最大エントロピー分布に及ぼす影響を、人為的にモーメントの値を変動させる感度分析から検討して、4次モーメントの変動に伴う最大エントロピー分布の感度は比較的鈍いことを示した。

6) 最後に、種々の時間単位の降雨・流量データのヒストグラムに対する最大エントロピー分布の適合度を検討して、2モーメント適合で十分な場合もあるが、一般的には4モーメント適合がよくヒストグラムを説明していることを明示した。

本章で得られた成果は、次章以下に述べる最大エントロピー分布の基礎をなすものであり、

その検討結果を十分に認識しておくことは極めて重要である。

参 考 文 献

- 1) 寒川典昭・荒木正夫：水文事象の頻度分析へのMEP導入について，土木学会論文報告集，第335号，pp.89～95，1987年7月。
- 2) 寒川典昭・荒木正夫・小林紀之：種々の水文量への1変数MEP分布の適用性，信州大学工学部紀要，第58号，pp.27～40，1985年9月。
- 3) Sonuga, J. O. : Principle of maximum entropy in hydrologic frequency analysis, Journal of Hydrology, Vol.17, pp.177～191, 1972.
- 4) Wragg, A. and Dowson, D. C. : Fitting continuous probability density function over $[0, \infty)$ using information theory ideas, IEEE Transactions on Information Theory, Vol.IT-16, pp.226～230, Mar., 1970.
- 5) 高棹琢馬・池淵周一：エントロピー的にみた降雨・流出変換特性とそのモデル化，京都大学防災研究所年報，第23号，B-2，pp.193～209，1980年4月。
- 6) Einbu, J. M. : On the existence of a class of maximum-entropy probability density function, IEEE Transactions on Information Theory, Vol.IT-23, pp.772～775, Nov., 1977.
- 7) 例えば，笠原芳郎：情報理論と通信方式，共立出版，p.150，1965年10月。
- 8) 例えば，前出5)。
- 9) 例えば，前出5)。
- 10) 例えば，前出5)。
- 11) Jowitt, P. W. : The extreme-value type-1 distribution and the principle of maximum entropy, Journal of Hydrology, Vol.42, pp.23～38, 1979.
- 12) Elderton, W. P. and Johnson, N. L. : Systems of frequency curves, Cambridge University Press, p.45, 1969.
- 13) 例えば，佐藤良一郎：数理統計学（増補版），培風館，pp.186～201，1943年5月。
- 14) 角屋 睦：水文統計論，土木学会水理委員会，p.2.，1964年7月。
- 15) 江藤剛治：降水量時系列の解析；土木学会誌，第63巻，第3号，78～83，1978年3月。

第3章 2変数最大エントロピー分布

第1節 概 説

2変数確率分布は、理論的にも、またその水文量への適用においても、確率分布による多変数統計解析の基礎となるため、本章では前章での議論を受けて、2変数最大エントロピー分布の定式化とその諸特性を検討する。

従来、水文統計における2変数分布として、指数分布^{1), 2), 3)}、正規分布⁴⁾、およびGamma分布^{1), 2)}、が用いられてきた。ところが、2変数指数分布、2変数正規分布は、周辺分布が指数分布、正規分布をする場合であり、このような水文データは特殊な場合に限られている。水文データの多くが、対数変換によって近似的に正規化できることから、実用的に2変数正規分布を用いる方法もあるが⁵⁾、これは水文量の母集団を推定するという立場からみると、さらに、研究を進展させる必要があろう。一方、2変数Gamma分布は、広く非対称形を表現することができ、長尾・角屋^{6), 7)}により応用面において有効な分布であることが報告されている。長尾らの研究は、2変数の水文統計に非対称性をとりこんだ点で、高く評価されるが、この分布は変形Bessel関数を含んでいるため、時には、その取り扱いに困難を生じる場合がある⁸⁾とともに、3変数以上に拡張して水文統計に応用することは、今のところ難かしい。

このような現状に鑑み、本章では、小標本を取り扱う水文統計にとって有効な分布であるとともに、密度関数式が簡潔で利用しやすいという特徴を有している2変数最大エントロピー分布を提案する^{9), 10)}。まず、第2節では、2変数最大エントロピー分布を定義し、そのパラメタ同定法を示すと同時に、この分布から、2変数の、一様、正規、相関のない指数分布を導出する。第3節では、仮定した母集団に対する適用性を、真のモーメントを使った場合について検討し、さらに、モーメントの安定性を把握するとともに、シミュレーションデータを使った場合の母集団推定とモーメントの変動に伴う推定分布の感度分析を行う。第4節では、いくつかの2地点で観測された水文データに対するこの分布の適用性を、情報として採用する統計モーメントの個数との関連で検討する。最後に、Appendixとして、第2節で示した最大エントロピー分布からの2変数正規分布と2変数指数分布の導出過程を詳述する。

なお、第3節、第4節において、適合度の改善および適合度の良否の判断は、3次元立体図、等確率線図、断面図、メッシュ確率を用いながら、必要に応じて複数の側面から検討していることを付記しておく。

第2節 確率密度関数の推定

2.1 定式化

ここでは、第2章2.1の定式化を、2つの確率変数 x, y の場合に拡張する。

いま、 x, y の同時確率密度関数を $p(x, y)$ とすると、式(2.1)のエントロピーは、次の相互エントロピーに置き換えられる。¹¹⁾

$$H(x, y) = - \int \int p(x, y) \ln p(x, y) dx dy \quad (3.1)$$

また、式(2.2)は確率密度関数が具備すべき条件、式(2.3)は情報の与え方を規定する式であったが、2変数の場合は、この2式は

$$\int \int p(x, y) dx dy = 1 \quad (3.2)$$

$$\int \int g_r(x, y) p(x, y) dx dy = E[g_r(x, y)], \quad r=1, 2, \dots, N \quad (3.3)$$

となる。ここに、 $g_r(\cdot)$ は任意関数である。

いま、 x, y の存在範囲を有限区間とした場合、式(3.2)を制約条件として式(3.1)を最大にする分布は、次節の式(3.8)で示すように2変数の一様分布となる。すなわち、この状態は、 x, y の生起に関して最もランダムであることを示している。一方、式(3.3)は、式(2.3)に対応した、 $p(x, y)$ に与える情報の表現式である。

したがって、式(3.2)、(3.3)を制約条件として式(3.1)を最大にする分布は、式(2.2)、(2.3)を制約条件として式(2.1)を最大にする分布と同じ立場をとることになる。このことから、前章2.1で記述した水文量の頻度解析に最大エントロピー分布を導入することの妥当性は、2変数の場合にも同様に成立することがわかる。また、上述の議論は3変数以上の多変数にも容易に拡張し得ることは言うまでもない。

そこで、上記の問題を、ここでも、ラグランジュの未定乗数法で解くことにする。ラグランジュ関数は

$$L=H(\mathbf{x}, \mathbf{y})+(\lambda_0-1)\{1-\int\int p(\mathbf{x}, \mathbf{y})d\mathbf{x}d\mathbf{y}\} \\ +\sum_{r=1}^N \lambda_r \{E[g_r(\mathbf{x}, \mathbf{y})]-\int\int g_r(\mathbf{x}, \mathbf{y})p(\mathbf{x}, \mathbf{y})d\mathbf{x}d\mathbf{y}\} \quad (3.4)$$

となるから，上式中の $p(\cdot)$ に関する変分を零とおくことにより

$$-\{1+\ln p(\mathbf{x}, \mathbf{y})\}-(\lambda_0-1)-\sum_{r=1}^N \lambda_r g_r(\mathbf{x}, \mathbf{y})=0 \quad (3.5)$$

が得られる。したがって，上式より

$$p(\mathbf{x}, \mathbf{y})=\exp\{-\lambda_0-\sum_{r=1}^N \lambda_r g_r(\mathbf{x}, \mathbf{y})\} \quad (3.6)$$

となる。ここに，ラグランジュ乗数 λ_r は推定分布のパラメタであること，その数は N 個であることは前章2.1で述べたとおりである。また，式(3.6)は，式(2.6)に対応した，一般的な2変数最大エントロピー分布であるから，ここでも， $g_r(\cdot)$ として種々の関数を与えることができる。

2.2 分布の導出

2変数最大エントロピー分布も，1変数の場合と同じように，形式的には無限個の分布形を記述することができる。ここでは主要な3つの既存分布の導出を考える。もちろん，この結果を用いて2変数最大エントロピー分布の位置付けを議論することができるが，これは第2章2.3と同様であること，および対比する2変数の一般分布系が存在しないことの原因から省略する。第4章，第5章においても同様の理由により，位置付けの議論は省略する。

1) 2変数一様分布

確率変数の存在範囲を $\mathbf{x}[K_1, K_2]$ ， $\mathbf{y}[L_1, L_2]$ として式(3.2)を次のように書き換える（式(3.3)は用いない）。

$$\int_{L_1}^{L_2} \int_{K_1}^{K_2} p(\mathbf{x}, \mathbf{y})d\mathbf{x}d\mathbf{y}=1, K_2>K_1, L_2>L_1 \quad (3.7)$$

このとき，式(3.6) は

$$P(x, y) = \frac{1}{(K_2 - K_1)(L_2 - L_1)} \quad (3.8)$$

となり，これは2変数一様分布である。

2) 2変数正規分布

$$g_1(x, y) = (x - \mu_x)^2 \quad (3.9)$$

$$g_2(x, y) = (y - \mu_y)^2 \quad (3.10)$$

$$g_3(x, y) = (x - \mu_x)(y - \mu_y) \quad (3.11)$$

とし，確率変数の存在範囲を $x(-\infty, \infty), y(-\infty, \infty)$ として，式(3.2), (3.3) を次のように書き換える。

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1 \quad (3.12)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x, y) dx dy = \sigma_x^2 \quad (3.13)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_y)^2 p(x, y) dx dy = \sigma_y^2 \quad (3.14)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) p(x, y) dx dy = \rho \sigma_x \sigma_y \quad (3.15)$$

ここに， μ_x ， μ_y ， σ_x^2 ， σ_y^2 ， ρ は定数である。

このとき，式(3.6) は

$$p(x, y) = \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{1/2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right\}\right] \quad (3.16)$$

となり、これは2変数正規分布である（この導出の詳細をAppendix Aに示す）。

3) 2変数指数分布（相関が零の場合）

$$g_1(x, y) = x \quad (3.17)$$

$$g_2(x, y) = y \quad (3.18)$$

とし、確率変数の存在範囲を $x[0, \infty), y[0, \infty)$ として、式(3.2), (3.3)を次のように書き換える。

$$\int_0^\infty \int_0^\infty p(x, y) dx dy = 1 \quad (3.19)$$

$$\int_0^\infty \int_0^\infty x p(x, y) dx dy = \eta_x \quad (3.20)$$

$$\int_0^\infty \int_0^\infty y p(x, y) dx dy = \eta_y \quad (3.21)$$

ここに、 η_x, η_y は定数である。

このとき、式(3.6)は

$$p(x, y) = \frac{1}{\eta_x \eta_y} \exp\left(-\frac{x}{\eta_x} - \frac{y}{\eta_y}\right) \quad (3.22)$$

となり、これは相関が零の場合の2変数指数分布である（この導出の詳細をAppendix Bに示す）。

2.3 パラメタ同定法

いま，1変数における4次までのモーメントは，それをもつ確率分布の形状と深くかかわりあっているため，ここでも，モーメントで情報を与えることを考える。 $x[0,\infty), y[0,\infty)$ とすると，式(3.2),(3.3)は次のように書き換えられる。

$$\int_0^{\infty} \int_0^{\infty} p(x,y) dx dy = 1 \quad (3.23)$$

$$\int_0^{\infty} \int_0^{\infty} x^a p(x,y) dx dy = x \mu_a, \quad a=1,2,\dots,N_a \quad (3.24)$$

$$\int_0^{\infty} \int_0^{\infty} y^b p(x,y) dx dy = y \mu_b, \quad b=1,2,\dots,N_b \quad (3.25)$$

$$\int_0^{\infty} \int_0^{\infty} x^c y^d p(x,y) dx dy = \mu_{cd}, \quad c=1,2,\dots,N_c, \quad d=1,2,\dots,N_d \quad (3.26)$$

ここに， $x \mu_a$ は x に関する原点まわりの a 次モーメント， $y \mu_b$ は y に関する原点まわりの b 次モーメント， μ_{cd} は x, y に関する原点まわりの c 次， d 次の相乗モーメントである。式(3.23)～(3.26)を制約条件とした2変数最大エントロピー分布は，式(3.6)より次式となる。

$$p(x,y) = \exp(-\alpha - \sum_{a=1}^{N_a} \beta_a x^a - \sum_{b=1}^{N_b} \gamma_b y^b - \sum_{c=1}^{N_c} \sum_{d=1}^{N_d} \delta_{cd} x^c y^d) \quad (3.27)$$

ここに， $\alpha, \beta_a, \gamma_b, \delta_{cd}$ は分布のパラメタであり，式(3.6)の $\lambda_0, \lambda_r (r=1,2,\dots,N)$ に相当するが，確率変数との関係を明示するためパラメタの表示を変更した。上記の分布を $2M(N_a, N_b, N_c, N_d)$ と略記する。

式(3.27)を式(3.24)～(3.26)に代入して，式(3.23)を用いて整理すると次式が得られる。

$$\begin{aligned}
& \int_0^\infty \int_0^\infty x^i \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy \\
&= x \mu_i \int_0^\infty \int_0^\infty \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy, \\
& \quad i=1, 2, \dots, Na
\end{aligned} \tag{3.28}$$

$$\begin{aligned}
& \int_0^\infty \int_0^\infty y^j \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy \\
&= y \mu_j \int_0^\infty \int_0^\infty \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy, \\
& \quad j=1, 2, \dots, Nb
\end{aligned} \tag{3.29}$$

$$\begin{aligned}
& \int_0^\infty \int_0^\infty x^k y^l \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy \\
&= \mu_{kl} \int_0^\infty \int_0^\infty \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy, \\
& \quad k=1, 2, \dots, Nc, \quad l=1, 2, \dots, Nd
\end{aligned} \tag{3.30}$$

したがって、式(3.27)のパラメタ $\beta_a, \gamma_b, \delta_{cd}$ は式(3.28)～(3.30)を解くことにより決定され、 α は式(3.27)を式(3.23)に代入して得られた次式から算定される。

$$\alpha = \ln \left\{ \int_0^\infty \int_0^\infty \exp\left(-\sum_{a=1}^{Na} \beta_a x^a - \sum_{b=1}^{Nb} \gamma_b y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd} x^c y^d\right) dx dy \right\} \tag{3.31}$$

ここで、式(3.28)～(3.30)は非線形であるため、 $\beta_a, \gamma_b, \delta_{cd}$ に関して解くために第2章2.4 で用いた方法を2変数の場合に拡張する。

いま、 $\beta_a, \gamma_b, \delta_{cd}$ の近似値として $\beta_a^0, \gamma_b^0, \delta_{cd}^0$ を与え、残差を $s \in a, r \in b, s \in cd$ とおくと次式が成立する。

$$\beta_a = \beta_a^0 + \beta_a \in a, \quad \gamma_b = \gamma_b^0 + \gamma_b \in b, \quad \delta_{cd} = \delta_{cd}^0 + \delta_{cd} \in cd \quad (3.32)$$

そこで，上式を式(3.28)～(3.30)に代入して $\beta_a \in a, \gamma_b \in b, \delta_{cd} \in cd$ に関して Taylor 展開し， $(\beta_a \in a)^2, (\gamma_b \in b)^2, (\delta_{cd} \in cd)^2$ 以上の微小量の項を無視して整理すると次式が得られる。

$$\sum_{a=1}^{Na} (A_{i+a,0} - \mu_i A_{a,0}) \beta_a \in a + \sum_{b=1}^{Nb} (A_{i,b} - \mu_i A_{0,b}) \gamma_b \in b + \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} (A_{i+c,d} - \mu_i A_{c,d}) \delta_{cd} \in cd = A_{i,0} - \mu_i A_{0,0}, \quad i=1,2,\dots,Na \quad (3.33)$$

$$\sum_{a=1}^{Na} (A_{a,j} - \mu_j A_{a,0}) \beta_a \in a + \sum_{b=1}^{Nb} (A_{0,j+b} - \mu_j A_{0,b}) \gamma_b \in b + \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} (A_{c,j+d} - \mu_j A_{c,d}) \delta_{cd} \in cd = A_{0,j} - \mu_j A_{0,0}, \quad j=1,2,\dots,Nb \quad (3.34)$$

$$\sum_{a=1}^{Na} (A_{k+a,l} - \mu_{kl} A_{a,0}) \beta_a \in a + \sum_{b=1}^{Nb} (A_{k,l+b} - \mu_{kl} A_{0,b}) \gamma_b \in b + \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} (A_{k+c,l+d} - \mu_{kl} A_{c,d}) \delta_{cd} \in cd = A_{k,l} - \mu_{kl} A_{0,0}, \quad k=1,2,\dots,Nc, \quad l=1,2,\dots,Nd \quad (3.35)$$

ここに

$$A_{i,j} = \int_0^\infty \int_0^\infty x^i y^j \exp\left(-\sum_{a=1}^{Na} \beta_a^0 x^a - \sum_{b=1}^{Nb} \gamma_b^0 y^b - \sum_{c=1}^{Nc} \sum_{d=1}^{Nd} \delta_{cd}^0 x^c y^d\right) dx dy, \quad i=1,2,\dots, \quad j=1,2,\dots \quad (3.36)$$

である。

式(3.33)～(3.35)は $\beta_a \in a, \gamma_b \in b, \delta_{cd} \in cd$ に関する $(N_a + N_b + N_c \times N_d)$ 元 1 次連立方程式であり，変数の数と方程式の数が一致するので解き得る。

やはり，ここでも，以下の手順に従って，式(3.33)～(3.35)を繰返し解くことにより，式(3.27)のパラメタ同定を行う。

1) 初期値 $\beta_a^0=0.0$ ， $\gamma_b^0=0.0$ ， $\delta_{cd}^0=0.0$ と仮定する。1変数の場合に計算した経験から判断して，このような初期値を設定した。以下の計算では，この初期値で十分により収束性を得ることができた。

2) 式(3.33)～(3.35)の連立方程式を $a \in a$ ， $r \in b$ ， $s \in cd$ に関して解く。

3) 得られた解を式(3.32)に代入して β_a ， γ_b ， δ_{cd} を計算し，これらを， β_a^0 ， γ_b^0 ， δ_{cd}^0 とする。

4) $a \in a$ ， $r \in b$ ， $s \in cd$ が十分小さくなるまで ($\leq 10^{-5}$)， 2)， 3) を繰返す。

5) $a \in a$ ， $r \in b$ ， $s \in cd$ が微小となったとき， β_a^0 ， γ_b^0 ， δ_{cd}^0 を推定パラメタとみなす。

6) 得られた推定パラメタを式(3.31)に代入して α を算定する。

ここで，収束判定基準を1変数の場合の 10^{-8} から 10^{-5} に変更したのは，式(3.36)の計算は2重の数値積分であるため，計算時間がかかなり増大すること，および多くの場合，パラメタの有効数字を 10^{-4} の桁までとれば実用的に十分な精度であることが1変数の場合から判断できたこと，の理由による。ただし，高精度が要求されるときは，この判定基準をさらに厳しくして計算している。

2.4 モーメントの基準化

ここでも，第2章 2.6で記述した理由により，統計モーメントの基準化を考える。確率変数 x ， y の平均が1となるような変数変換，すなわち

$$u=x/\ x\ \mu_1\ ,\ v=y/\ y\ \mu_1 \quad (3.37)$$

を行うと， u ， v に関する確率密度関数 $q(u,v)$ は，Jacobian $J(\cdot)$ を介して

$$q(u,v)=J\left(\frac{x,y}{u,v}\right)p(x,y) \quad (3.38)$$

となる。ここに

$$J\left(\frac{\mathbf{x}, \mathbf{y}}{\mathbf{u}, \mathbf{v}}\right) = \mathbf{x} \mu_1 \cdot \mathbf{y} \mu_1 \quad (3.39)$$

であるから、式 (3.38) は

$$q(\mathbf{u}, \mathbf{v}) = p(\mathbf{x}, \mathbf{y}) \cdot \mathbf{x} \mu_1 \cdot \mathbf{y} \mu_1 \quad (3.40)$$

あるいは

$$p(\mathbf{x}, \mathbf{y}) = q(\mathbf{u}, \mathbf{v}) / (\mathbf{x} \mu_1 \cdot \mathbf{y} \mu_1) \quad (3.41)$$

と表現できる。

上述の変数変換を行うと、式(2.23)～(2.26)は次のように書き換えられる。

$$\int_0^\infty \int_0^\infty q(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} = 1 \quad (3.42)$$

$$\int_0^\infty \int_0^\infty u^a q(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} = \mathbf{x} \mu_a / (\mathbf{x} \mu_1)^a \equiv \mathbf{x} m_a, \quad a=1, 2, \dots, N_a \quad (3.43)$$

$$\int_0^\infty \int_0^\infty v^b q(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} = \mathbf{y} \mu_b / (\mathbf{y} \mu_1)^b \equiv \mathbf{y} m_b, \quad b=1, 2, \dots, N_b \quad (3.44)$$

$$\int_0^\infty \int_0^\infty u^c v^d q(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} = \mu_{cd} / \{(\mathbf{x} \mu_1)^c (\mathbf{y} \mu_1)^d\} \equiv m_{cd} \quad (3.45)$$

$$c=1, 2, \dots, N_c, \quad d=1, 2, \dots, N_d$$

ここに、 $\mathbf{x} m_0 = \mathbf{x} m_1 = \mathbf{y} m_0 = \mathbf{y} m_1 = m_{00} = m_{01} = m_{10} = m_{11} = 1$ であるため、式(3.43), (3.44), (3.45)で定義されたモーメントをそれぞれ、 \mathbf{x} に関する基準化されたモーメント、 \mathbf{y} に関する基準化されたモーメント、および基準化された相乗モーメントと呼ぶことにする。

次に、 $p(\cdot)$ と $q(\cdot)$ のエントロピーの関係を求めると、一変数の場合と同様に簡単な式展開から

$$H(x, y) = H(u, v) + \ln(x \mu_1 \cdot y \mu_1) \quad (3.46)$$

が導出される。従って、第2章2.6の議論を受けて、ここでも、実際の適用では、式(3.42)～(3.45)を制約条件として $H(u, v)$ を最大にする $q(\cdot)$ を求めておき、しかる後に、式(3.41)により $p(\cdot)$ に変換すればよい。

第3節 既存分布を母集団とした場合の適用性の検討

3.1 真のモーメントを用いた場合の適用性

3.1.1 仮定した母集団

第2章3.1.1で述べた理由により、ここでも母集団を仮定する必要がある。母集団としては非対称形を表現できること、多くの形状を取り得ること、の条件を満たすことが望ましいため、本節では2変数Gamma分布を取り上げる。

周辺分布が形状母数 $\nu_1 = n+m$, $\nu_2 = n$ ($m, n \geq 0$)のGamma分布である2変数 x, y についての密度関数は、井沢²⁾により次のように書かれている。

$$p(x, y) = \frac{(xy)^{(n-1)/2} x^m}{\Gamma(n) \Gamma(m) (\sigma_1 \sigma_2)^{(n+1)/2} \sigma_1^m (1-\rho) \rho^{(n-1)/2}} \cdot \exp\left\{-\frac{x}{\sigma_1(1-\rho)} - \frac{y}{\sigma_2(1-\rho)}\right\} \int_0^1 (1-t)^{(n-1)/2} t^{m-1} \exp\left\{-\frac{\rho xt}{\sigma_1(1-\rho)}\right\} \cdot I_{n-1}\left[\frac{2\rho^{1/2}}{1-\rho} \{(xy(1-t)/(\sigma_1 \sigma_2))^{1/2}\} \right] dt \quad (3.47)$$

ここに、 $\Gamma(\cdot)$ はGamma関数、 $I_\nu(\cdot)$ は ν 次の変形Bessel関数であり、 σ_1, σ_2, ρ は定数である。以下では、形状母数 ν_1, ν_2 が $\nu_1 \neq \nu_2$ のとき式(3.47)を $2G(\nu_1, \nu_2, \rho)$ 、 $\nu_1 = \nu_2 = \nu$ のとき $2G(\nu, \rho)$ と略記し、母集団と仮定する。このとき、尺度母数 σ_1, σ_2 は、式(3.45)の $m_{10} = m_{01} = 1$ となるように決定される。

一方、2変数 x, y についての、式(3.47)の原点まわりの p, q 次のモーメント μ_{pq} は次式で与えられている¹⁾。

$$\mu_{cd} = \frac{\Gamma(n+m+c) \Gamma(n+d)(1-\rho)^{n+d} \sigma_1^c \sigma_2^d}{\{\Gamma(n)\}^2 \Gamma(m)} I(c,d) \quad (3.48)$$

ここに

$$I(c,d) = \int_0^1 z^{n-1} (1-z)^{m-1} (1-\rho z)^{-n-m-d} R(c,d) dz \quad (3.49)$$

であり

$$R(c,d) = F(-m-c, -d; n; \rho z) \quad (3.50)$$

である。ただし、 $F(\cdot)$ は超幾何関数を表現している。具体的な c, d を与えた場合の式(3.48)の計算結果は、2次以下のモーメントについては文献7)に記載されており、 μ_{60} 、 μ_{06} 、 μ_{22} ま

表3.1 2G(1,0.2) , 2G(2,0.5) および2G(4,0.5) を母集団とした場合の最大エントロピー分布のパラメタの値

2G(1,0.2)	2M(3,3,1,1)	2M(4,4,1,1)
α	-0.14418	-0.13749
$\beta_1 (= \gamma_1)$	1.1441	1.1308
$\beta_2 (= \gamma_2)$	-0.0029132	0.0069515
$\beta_3 (= \gamma_3)$	0.0026707	0.00050528
$\beta_4 (= \gamma_4)$	—	0.00013316
δ_{11}	-0.15045	-0.15042

2G(2,0.5)	2M(2,2,1,1)	2M(4,4,1,1)	2M(4,4,2,2)
α	0.57084	1.6852	1.2734
$\beta_1 (= \gamma_1)$	0.14688	-2.0921	-1.5278
$\beta_2 (= \gamma_2)$	0.77881	2.8953	3.2772
$\beta_3 (= \gamma_3)$	—	-0.70465	-0.89304
$\beta_4 (= \gamma_4)$	—	0.068628	0.077565
δ_{11}	-1.1866	-1.0480	-2.9656
$\delta_{12} (= \delta_{21})$	—	—	0.40641
δ_{22}	—	—	-0.067129

2G(4,0.5)	2M(2,2,1,1)	2M(4,4,1,1)	2M(4,4,2,2)
α	2.1283	5.3627	4.5707
$\beta_1 (= \gamma_1)$	-1.8801	-8.4525	-7.2592
$\beta_2 (= \gamma_2)$	2.3276	9.7708	10.475
$\beta_3 (= \gamma_3)$	—	-3.1215	-3.6144
$\beta_4 (= \gamma_4)$	—	0.39584	0.42859
δ_{11}	-2.6116	-2.3235	-6.3267
$\delta_{12} (= \delta_{21})$	—	—	1.0551
δ_{22}	—	—	-0.22084

で拡張したモーメントについても容易に求められる。

3.1.2 等しい形状母数を持つ母集団の場合

ここでは、指数分布 $2G(1,0.2)$ ，緩やかな分布 $2G(2,0.5)$ ，鋭くとがった分布 $2G(4,0.5)$ を母集団として採用する。表3.1は，それぞれの母集団がもつ基準化されたモーメントを制約条件として得られた式(3.27)の最大エントロピー分布のパラメタの値である。適合度は，3次元立体図と等確率線図を用いて判定する（全体図比較）とともに，後述の他の2側面からも検討を加え，この判定の妥当性を検証する。

図3.1は，仮定した母集団 $2G(1,0.2)$ と，最大エントロピー分布 $2M(3,3,1,1)$ ， $2M(4,4,1,1)$ の内 $2M(3,3,1,1)$ の場合の3次元立体図である。3者は，全体的にはもちろんのこと， u ， v の小さい領域および大きい領域においても，ほぼ一致しており，このような指数分布に対しては $2M(3,3,1,1)$ で十分よい再現が得られることがわかる。なお，3.1.2，3.1.3，3.2.2，3.2.3の3次元立体図には，等確率線図を併記して適合性を判断しているが，等確率線図の掲載は省略した。

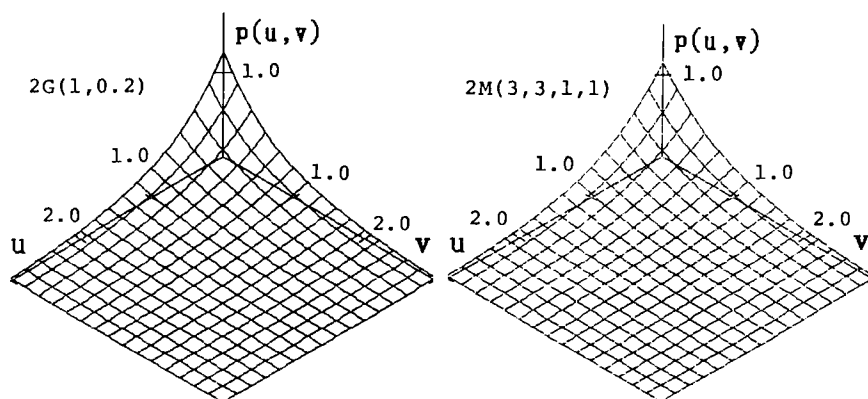


図3.1 $2G(1,0.2)$ と $2M(3,3,1,1)$ の3次元立体図

図3.2は，仮定した母集団 $2G(2,0.5)$ と，最大エントロピー分布 $2M(2,2,1,1)$ ， $2M(4,4,1,1)$ ， $2M(4,4,2,2)$ の内 $2M(4,4,2,2)$ の場合の3次元立体図である。 $2G(2,0.5)$ のモードの位置はかなり原点の近くにある。 $2M(2,2,1,1)$ の適合度は $2G(2,0.5)$ よりフラットになり，あまりよくない。 $2M(4,4,1,1)$ は母集団の形状を概略的に表現しているが，まだ十分な適合度は得られていない。 $2M(4,4,2,2)$ になると u 軸および v 軸周辺に適合度の改善がみられるが， $2G(1,0.2)$ の場合に比べれば適合度は劣る。しかし，極値水文統計上重要な u ， v の大きな領域では母集団に対して十分よく適合している。

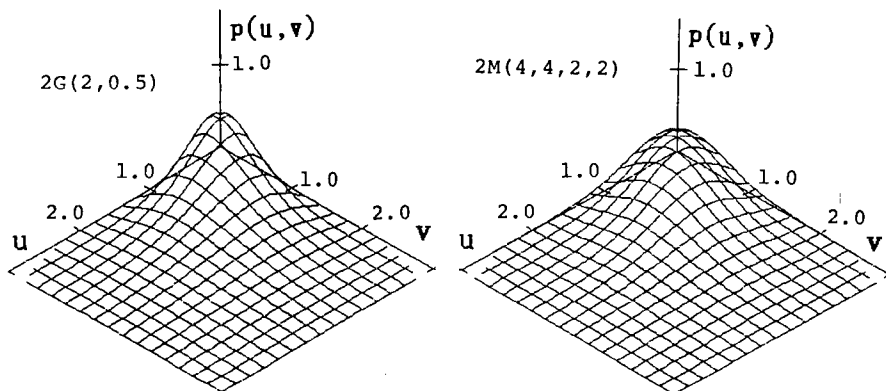


図3.2 $2G(2, 0.5)$ と $2M(4, 4, 2, 2)$ の3次元立体図

図3.3 は、仮定した母集団 $2G(4, 0.5)$ と、最大エントロピー分布 $2M(2, 2, 1, 1)$, $2M(4, 4, 1, 1)$, $2M(4, 4, 2, 2)$ の内 $2M(4, 4, 2, 2)$ の場合の3次元立体図である。 $2G(4, 0.5)$ は $2G(2, 0.5)$ に比べてモード付近での密度関数の値が大きく、その位置も原点から離れている。 $2M(2, 2, 1, 1)$ は $2G(4, 0.5)$ よりフラットになり、概略的にしか母集団をとらえていないが、 $2G(2, 0.5)$ の場合 $2M(2, 2, 1, 1)$ に比べればよい適合度を示している。 $2M(4, 4, 1, 1)$ は $2M(2, 2, 1, 1)$ に比べてかなり母集団に対する適合度を増し、モード付近での密度関数の値、モードの位置がほぼ母集団と一致している。 $2M(4, 4, 2, 2)$ にすると、 u , v の小さい領域では $v=u$ 軸に沿う方向に収縮され、大きい領域では拡大されて、それらの領域でより母集団との適合度が改善されてくる。

次に、特に u , v の小さい領域、および大きい領域に注目してみるために、 $v=0$ ($2G(1, 0.2)$ の場合のみ)、 0.5 , 1.0 , 1.5 , 2.0 における u 方向の断面図を普通紙上に、 $v=0$ ($2G(1, 0.2)$ の

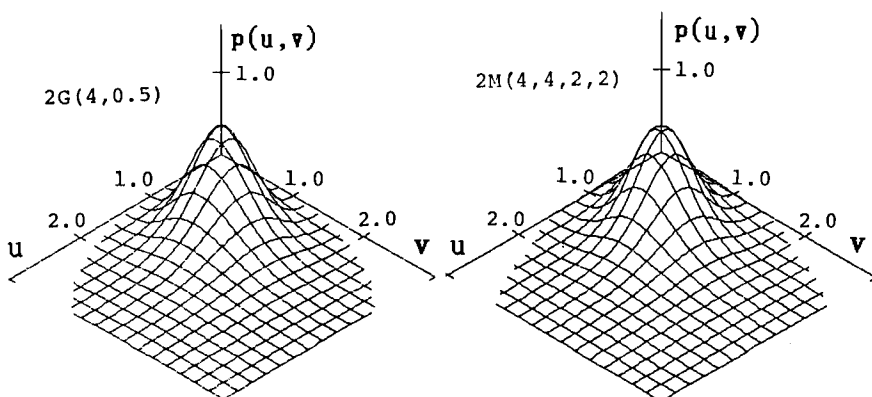


図3.3 $2G(4, 0.5)$ と $2M(4, 4, 2, 2)$ の3次元立体図

場合のみ) , 0.5, 1.0, 1.5, 2.0, 3.0, 4.0におけるu方向の断面図を片対数紙上に描いて適合度を検討した(断面図比較)。さらに, 図3.1 ~ 3.3 の立体図を0.5 区間のメッシュに切りメッシュ上の体積をメッシュ面積で除した値により適合度比較を行った(メッシュ確率比較)。これらの比較による適合度の結果は, 前述の考察を支持するものであった。

3.1.3 異なる形状母数を持つ母集団の場合

ここでは, 一方の変数の周辺分布が指数分布, 他方は山形分布をする $2G(1,2,0.2)$, 両者とも同一でない山形分布をする $2G(4,2,0.5)$ を母集団として採用する。表3.2 は, それぞれの母集団がもつ基準化されたモーメントを制約条件として得られた式(3.27)の最大エントロピー分布のパラメタの値である。適合度の判定は3.1.2 の全体図比較を用いて行う。

表3.2 $2G(2,1,0.2)$, $2G(4,2,0.5)$ を母集団とした場合の最大エントロピー分布のパラメタの値

$2G(2,1,0.2)$	$2M(2,2,1,1)$	$2M(4,4,1,1)$	$2M(4,4,2,2)$
α	0.44606	1.0496	0.98926
β_1	-0.22343	-2.7096	-2.6227
β_2	0.47515	2.8838	2.8747
β_3	—	-0.76278	-0.76496
β_4	—	0.070337	0.069914
γ_1	1.1615	1.1564	1.2405
γ_2	0.010136	0.0050997	0.0011730
γ_3	—	0.00046649	-0.00068954
γ_4	—	0.000013002	0.0000043483
δ_{11}	-0.18365	-0.16952	-0.30683
δ_{12}	—	—	0.013660
δ_{21}	—	—	0.029066
δ_{22}	—	—	-0.0016040

$2G(4,2,0.5)$	$2M(2,2,1,1)$	$2M(4,4,1,1)$	$2M(4,4,2,2)$
α	1.6458	3.9356	3.4944
β_1	-2.6333	-9.4588	-8.8829
β_2	1.9494	9.8161	9.9841
β_3	—	-3.2788	-3.4237
β_4	—	0.41419	0.41049
γ_1	0.45067	-2.0656	-1.2971
γ_2	0.59622	2.9234	2.9518
γ_3	—	-0.74772	-0.79726
γ_4	—	0.070038	0.072411
δ_{11}	-1.1016	-0.95624	-2.5731
δ_{12}	—	—	0.24403
δ_{21}	—	—	0.51503
δ_{22}	—	—	-0.066436

図3.4 は, 仮定した母集団 $2G(2,1,0.2)$ と, 最大エントロピー分布 $2M(2,2,1,1)$, $2M(4,4,1,1)$, $2M(4,4,2,2)$ の内 $2M(4,4,1,1)$ の場合の3次元立体図である。 $2M(2,2,1,1)$ の適合度は, $2G(2,1,0.2)$ よりフラットになり, あまりよくない。 $2M(4,4,1,1)$ にするとv軸上付近では母集団と十分一致していない部分があるが, 全体的にはかなり適合度が改善され, u, vの大き

い領域では，十分な精度で母集団に合致している。 $2M(4,4,2,2)$ は $2M(4,4,1,1)$ とほぼ一致し，この場合相乗モーメントの次数の増加の効果はわずかであることがわかる。

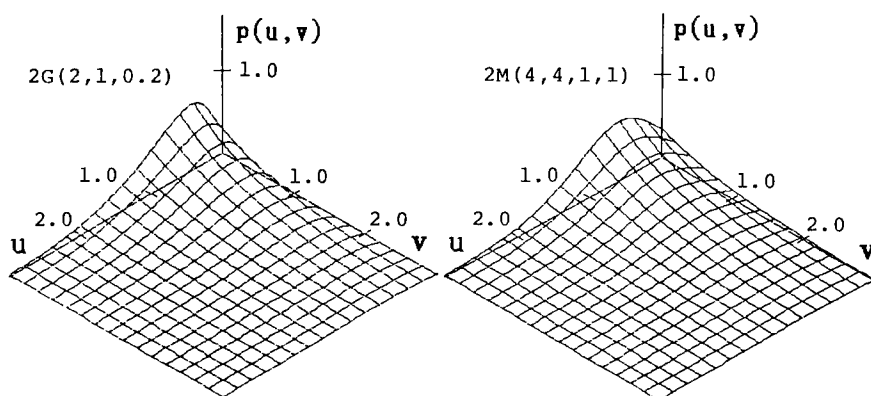


図3.4 $2G(2,1,0.2)$ と $2M(4,4,1,1)$ の3次元立体図

図3.5 は，仮定した母集団 $2G(4,2,0.5)$ と，最大エントロピー分布 $2M(2,2,1,1)$, $2M(4,4,1,1)$, $2M(4,4,2,2)$ の内 $2M(4,4,2,2)$ の場合の3次元立体図である。 $2G(4,2,0.5)$ の u 方向の分散は v 方向の分散より小さくなっている。この場合も $2M(2,2,1,1)$ は $2G(4,2,0.5)$ よりフラットになり，あまり適合度はよくない。 $2M(4,4,1,1)$ にするとかなり適合度が改善され， $2M(4,4,2,2)$ にするとさらに u 軸付近での改善がみられる。部分的にみると， $2M(4,4,2,2)$ は u 軸上付近でまだ改善の余地が残されているが， u ， v の大きい領域ではよく母集団に合致している。

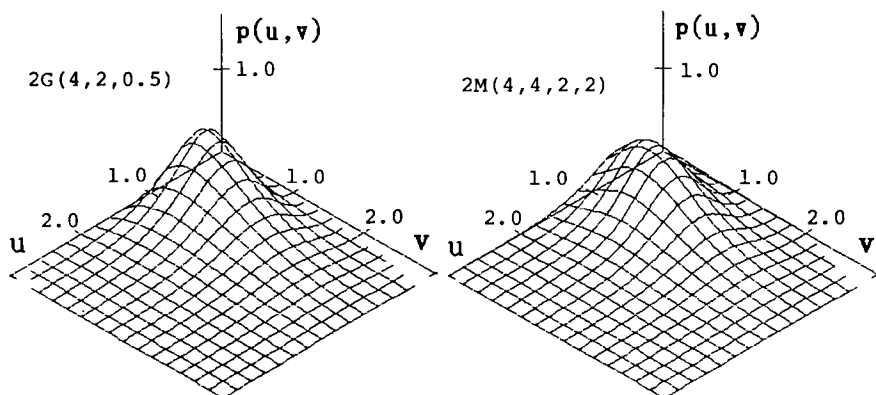


図3.5 $2G(4,2,0.5)$ と $2M(4,4,2,2)$ の3次元立体図

3.2 モーメントの安定性と感度分析

3.2.1 モーメントの安定性

ここでは、モーメントの持つ誤差を検討するために、 $2G(1,0.2)$, $2G(2,0.5)$, $2G(4,0.5)$ を母集団として個数 n ($n=10, 20, \dots, 150$) のデータを50回発生させ、 $m_{10}, m_{20}, m_{30}, m_{40}, m_{50}, m_{60}, m_{11}, m_{12}, m_{21}, m_{22}$ の平均と標準偏差を求めた。図3.6 は、1例として母集団を $2G(2,0.5)$ とした場合の m_{40} と m_{60} の安定性を示したものである。全体的にみて、データ数が60個ないし70個付近から平均値が母集団のモーメントに近づき、標準偏差が小さくなる様子がうかがわれる。また、高次モーメントになるにしたがい標準偏差が大きくなり、前項の適合性との関連から、最大エントロピー分布への情報としては4次モーメントまでに留めることが妥当と判断される。

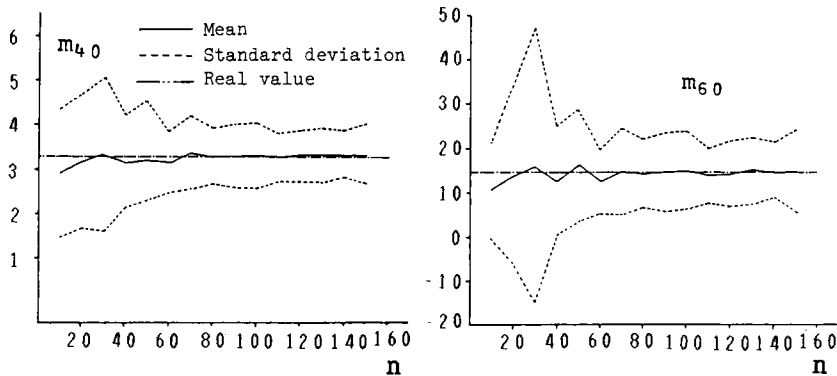


図3.6 基準化されたモーメント m_{40} と m_{60} の安定性
(母集団: $2G(4,0.5)$)

3.2.2 データ数の増加と母集団推定

ここでは、母集団から発生させるデータを $n=10, 20, \dots, 150$ と増加させ、検討を要するデータ数に対して最大エントロピー分布を求めた。

1) $2G(1,0.2)$ と $2N(3,3,1,1)$

母集団を $2G(1,0.2)$ とし、 $n=20, 40, \dots, 140$ ごとに $2N(3,3,1,1)$ を求め、3次元立体図と等確率線図を描いた。図3.7 は $n=40$ の場合の3次元立体図であり、母集団分布と比較して、原点付近はすこし高くなっているが、全体的にはよく一致しているとともに、 u, v の大きい領域ではよい適合度を示している。また、推定分布の形状は $n=40$ から母集団に近くなるが、その後のデータ数になっても、シミュレートした乱数の特性により、モーメントのサンプリングエラーが大きくなることもあり、 $n=40$ の分布より適合度が劣る場合もあった。しかし、平均的にみると、データ数の増加とともに、母集団の分布に近い推定分布が得られることがわかった。

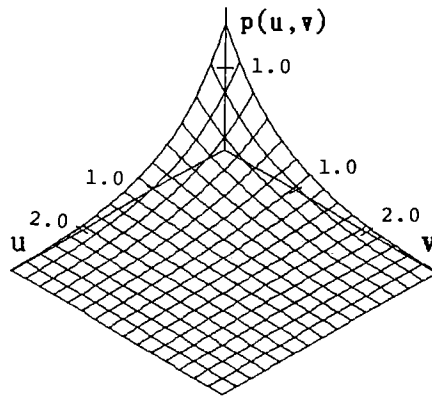


図3.7 $n = 40$ の場合の $2M(3,3,1,1)$ の 3 次元立体図
(母集団: $2G(1,0.2)$)

2) $2G(4,0.5)$ と $2M(2,2,1,1), 2M(4,4,2,2)$

母集団を $2G(4,0.5)$ とし, $n=10, 20, \dots, 150$ に対する $2M(2,2,1,1)$ と, $n=70, 90, 100, 110, 130, 140, 150$ の $2M(4,4,2,2)$ を求めた。 $n=10, 20, 30, 40, 50, 60, 80, 120$ の $2M(4,4,2,2)$ は求まらなかった。その理由は, これらのデータ数で計算された統計モーメントが, $2M(4,4,2,2)$ の存在条件を満足しなかったためと推察される。 $2M(2,2,1,1)$ の場合, 20個ごとに 3 次元立体図と等確率線図を描いた結果, $n=60$ 程度から真値のモーメントを用いた $2M(2,2,1,1)$ に, 全体的にも, u, v の小さい領域, 大きい領域においてもよく近似していることがわかった。 $2M(4,4,2,2)$ の場合, 求まった 7 ケースとも 3 次元立体図と等確率線図を描いた結果, どのケースも真のモーメントを用いた $2M(4,4,2,2)$ に, 上述と同様によく近似していることがわかった。図

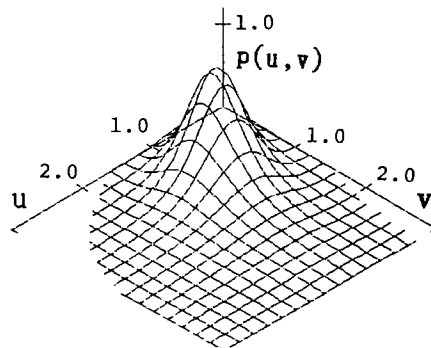


図3.8 $n = 70$ の場合の $2M(4,4,2,2)$ の 3 次元立体図
(母集団: $2G(4,0.5)$)

3.8 は $n=70$ の場合の $2M(4,4,2,2)$ の3次元立体図である。

3.2.3 同数の異種データによる母集団推定

ここでは、母集団を $2G(4,0.5)$ とし、1組70個のデータを5組発生させて、各ケースごとに $2M(2,2,1,1)$ と $2M(4,4,2,2)$ を求めた。図3.9 は、真値のモーメントを用いた場合と比較して最も離れた形状を示したケース3の $2M(2,2,1,1)$ と $2M(4,4,2,2)$ の3次元立体図である。ケースごとの変動は、データ数が少ない程大きくなることが予想されるが、70個のデータ数になってもこの程度の変動が存在することを認識しておかねばならない。また、 $2M(2,2,1,1)$ の方が、 $2M(4,4,2,2)$ より大きな変動を示すことも明らかになった。

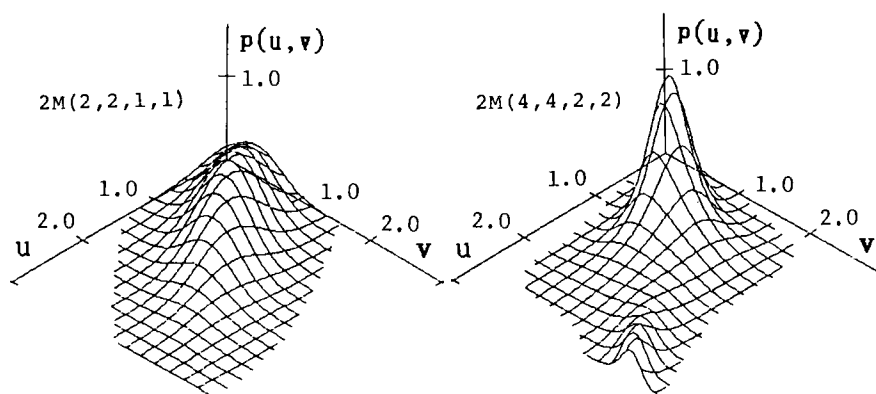


図3.9 5 ケースの内、最も適合度が悪い場合の $2M(2,2,1,1)$ と $2M(4,4,2,2)$ の3次元立体図（母集団： $2G(4,0.5)$ ）

3.2.4 モーメントの変動による感度分析

ここでは、モーメントの人為的変動が最大エントロピー分布に及ぼす影響を分析する。母集団として $2G(1,0.2)$, $2G(4,0.5)$ を採用し、変動させるモーメントは最高次数の周辺分布に関するモーメントと最高次数の相乗モーメントとし、それ以外はすべて真値に固定した。

このとき、採用した最大エントロピー分布が、3次元立体図から判断して真のモーメントを使った最大エントロピー分布に形状的に近似していた区間は次のとおりであった。

$2G(1,0.2)$ に対する $2M(3,3,1,1)$

$$6.0 \leq m_{30} (\text{真値}=6.0) \leq 10.0, \quad 1.01 \leq m_{11} (1.2) \leq 1.75 \quad (3.51)$$

$2G(1,0.2)$ に対する $2M(4,4,1,1)$

$$23.0 \leq m_{40}(24.0) \leq 24.39, \quad 0.8 \leq m_{11}(1.2) \leq 1.3 \quad (3.52)$$

2G(4,0.5) に対する 2M(2,2,1,1)

$$1.20 \leq m_{20}(1.25) \leq 1.35, \quad 1.05 \leq m_{11}(1.125) \leq 1.175 \quad (3.53)$$

2G(4,0.5) に対する 2M(4,4,2,2)

$$3.22 \leq m_{40}(3.2813) \leq 3.795, \quad 2.35 \leq m_{22}(2.3823) \leq 2.43 \quad (3.54)$$

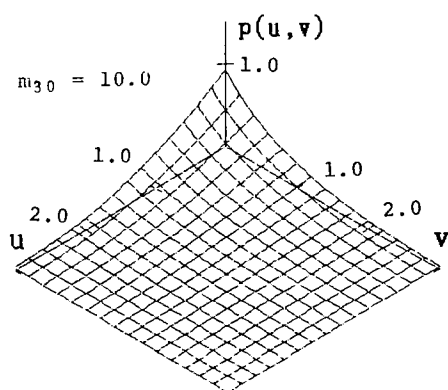


図3.10 式(3.51)の m_{30} の上限値での 2M(3,3,1,1) の 3次元立体図 (母集団: 2G(1,0.2))

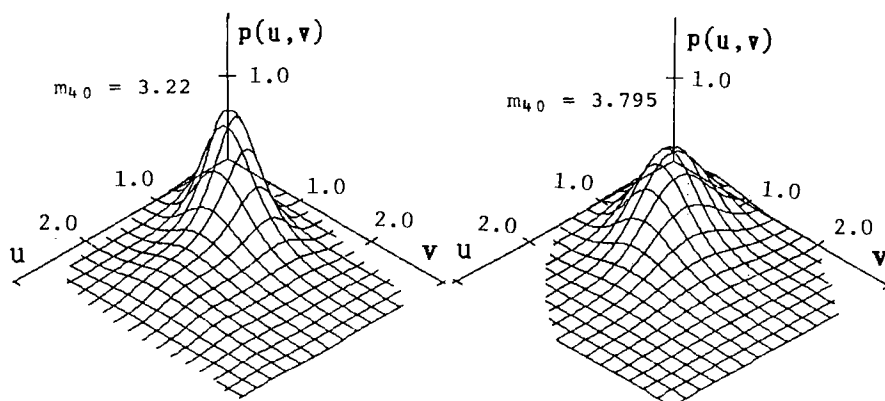


図3.11 式(3.54)の m_{40} の上, 下限値での 2M(4,4,2,2) の 3次元立体図 (母集団: 2G(4,0.5))

図3.10は、式(3.51)で示した m_{30} の上限値での $2M(3,3,1,1)$ である。ただし、 m_{30} の下限値での $2M(3,3,1,1)$ は、真値のモーメントを用いた場合であり、図3.1 に掲載している。また、図3.11は、式(3.54)で示した m_{40} の上・下限値での $2M(4,4,2,2)$ の3次元立体図である。

第4節 実データへの適用

4.1 2地点流量の場合

石原・長尾⁵⁾は、合流河川における治水計画に2変数正規分布を適用した。ここでは、同じ目的のために、2変数最大エントロピー分布を求める。

使用したデータは、図3.12の立ヶ花地点29年間の年最大流量に対応する小市、杭瀬下の最大流量である。ここでは、両者の値を各々の平均で除した基準化した値をそれぞれ確率変数 u 、 v として $2M(2,2,1,1)$ と $2M(4,4,1,1)$ を求めた。このときのパラメタの値を表3.3 に、 $2M(4,4,1,1)$ の3次元立体図を図3.13に、 $2M(2,2,1,1)$ 、 $2M(4,4,1,1)$ 、およびデータのメッシュ確率を図3.14に示している。データ数が少ないので厳密な議論はできないが、 $2M(2,2,1,1)$

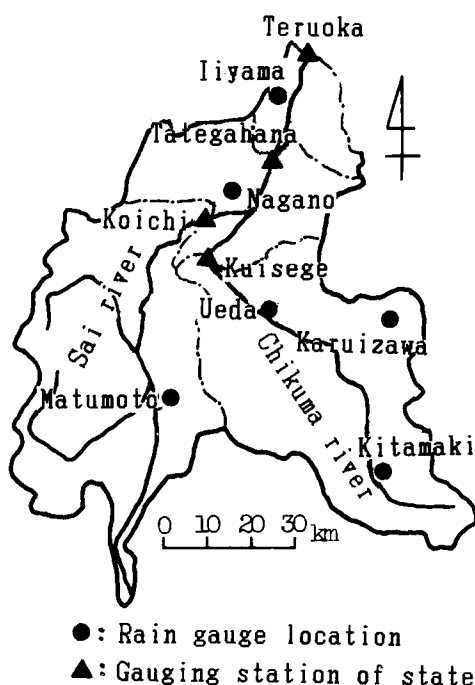


図3.12 千曲川流域図（長野県）

から $2M(4,4,1,1)$ にするとデータに対する適合度が改善され、 u あるいは v の大きいところに存在するデータも説明できるように、分布形状が移行している様子がうかがわれる。

表3.3 2地点流量データから求めた最大エントロピー分布のパラメタの値

	$2M(2,2,1,1)$	$2M(4,4,1,1)$
α	1.8875	9.5384
β_1	-4.1926	-33.053
β_2	2.6726	39.458
β_3		-18.184
β_4		2.9935
γ_1	1.4446	-1.9346
γ_2	0.22082	5.3094
γ_3		-2.5014
γ_4		0.37225
δ_{11}	-1.0654	-0.95042

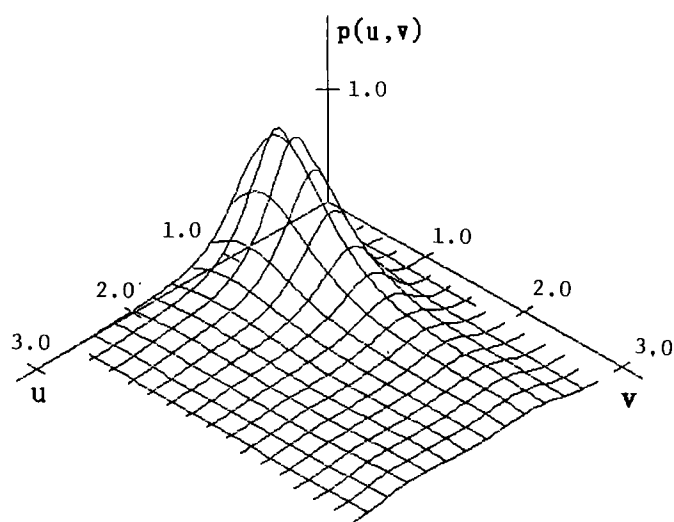


図3.13 2地点流量データより求めた $2M(4,4,1,1)$ の3次元立体図

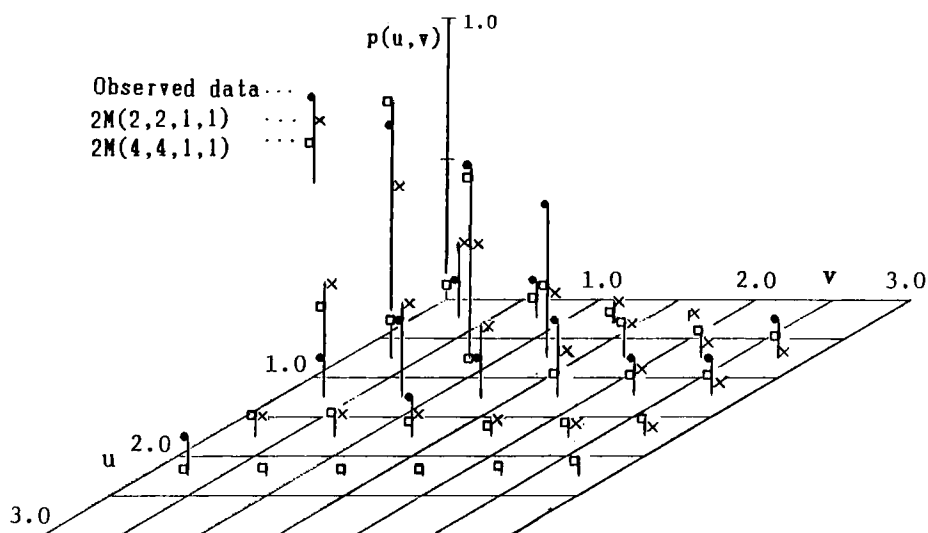


図3.14 $2M(2,2,1,1)$, $2M(4,4,1,1)$ およびデータのメッシュ確率

4.2 年最大日降雨と対応日流量の場合

ここでは、図3.12に示した上田の28年間の年最大日降雨と、それに対応する杭瀬下の日流量の値を、各々の平均で除して基準化し、それらを確率変数 u , v として最大エントロピー分布を求めた。表3.4 は $2M(2,2,1,1)$ のパラメタの値である。この場合、モーメントの最大次数を3以上に増加させた最大エントロピー分布は求まらなかった。図3.15は $2M(2,2,1,1)$ の3次元立体図であり、図3.16は $2M(2,2,1,1)$ とデータのメッシュ確率を比較したものである。 $2M(2,2,1,1)$ は、データの全体的傾向をとらえているが、両者のモードの位置が異っており、データは必ずしも推定分布のように v 軸方向に減少傾向を示していない等、十分な適合度をもつ

表3.4 降雨－流量データから求めた最大エントロピー分布のパラメタの値

	$2M(2,2,1,1)$
α	3.9403
β_1	-9.1643
β_2	5.4547
γ_1	1.6623
γ_2	0.37273
δ_{11}	-1.7157

とは言えない。また、4.1 も同様であるが、データ数が少ないため、データ1個に付与される確率が大きくなり、分布形の裾の部分の適合度は考察しにくい。

4.1 , 4.2 より、実データに対しても4次モーメントまで情報とすることが望まれる。

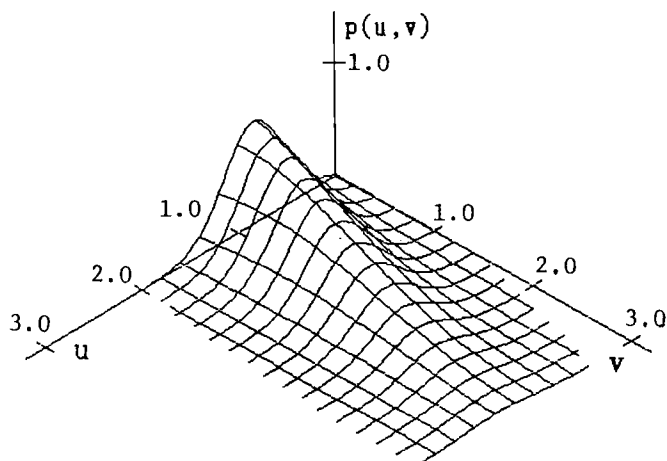


図3.15 降雨-流量データより求めた
 $2M(2,2,1,1)$ の3次元立体図

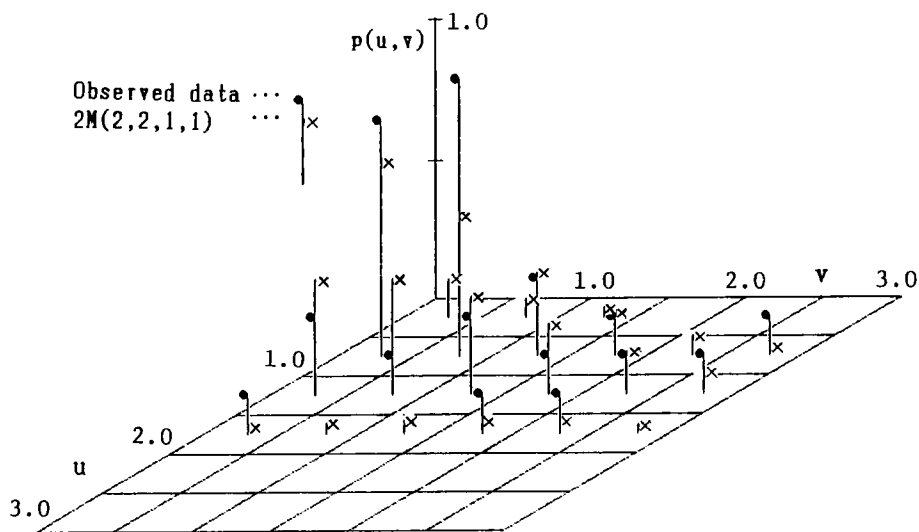


図3.16 $2M(2,2,1,1)$ とデータのメッシュ確率

第5節 要 約

本章では、2変数最大エントロピー分布を提案した。さらに、この分布は、多変数分布の基礎となるため、その持つ特性を種々の側面から検討した。得られた成果を要約すると次のようになる。

1) 2変数最大エントロピー分布を定式化した。また、1変数の場合を拡張して、そのパラメタ同定法を示した。

2) 2変数最大エントロピー分布から、2変数一様、正規、相関のない指数分布を導出した。特に、2変数正規分布と2変数指数分布の導出については、Appendixで詳述した。

3) いくつかの典型的な形状を示す2変数ガンマ分布を母集団と仮定して、その真のモーメントを情報とした最大エントロピー分布で母集団推定を行った。その結果から、母集団が指数分布をする場合は3次モーメントまで、その他の場合は4次モーメントまで情報とすれば、十分によい再現が得られることを示した。

4) 統計モーメントの安定性について、シミュレーションデータを用いて調べた。ついで、3)の検討を踏まえて、統計モーメントの推定誤差が最大エントロピー分布に及ぼす影響を、シミュレーションデータと人為的にモーメントを変動させる感度分析の立場から検討した。その結果を母集団分布と推定分布の組み合わせごとに、シミュレーションデータの場合はデータ数で、人為的にモーメントを変動させた場合はモーメントの変動域で、モーメントの推定誤差の許容範囲として提示した。

5) 相関のある2地点で観測された水文データに対して、2変数最大エントロピー分布の適合性を検討した。この場合も、3)と同様、4次モーメントまで情報とすればよい適合性が得られることを示した。

以上の結果から、2変数最大エントロピー分布は、それがもつ定式化における思想的な背景をも含めて、水文統計を行う上で有効な分布であることが明らかになった。当然、本章で詳述した特性は、3変数以上へ拡張した場合にも受けつがれるものであり、ここで述べた事項は、次章の多変数最大エントロピー分布を論ずる上で、多くの示唆に富む内容を含んでいる。

Appendix A 2変数最大エントロピー分布からの2変数正規分布の導出

$x(-\infty, \infty)$, $y(-\infty, \infty)$ とし、式(3.12)～(3.15)を制約条件として式(3.1)を最大にする $p(\cdot)$ は、式(3.16)の2変数正規分布となることを示す。

この問題のラグランジュ関数は

$$\begin{aligned}
L = & H(x, y) + (\alpha - 1) \left\{ 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy \right\} + \beta \left\{ \sigma_x^2 - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x, y) dx dy \right\} \\
& + \gamma \left\{ \sigma_y^2 - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_y)^2 p(x, y) dx dy \right\} \\
& + \delta \left\{ \rho \sigma_x \sigma_y - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) p(x, y) dx dy \right\}
\end{aligned} \quad (1)$$

となる。上式を $p(\cdot)$ について変分をとり零とおくと、

$$- \{1 + \ln p(x, y)\} - (\alpha - 1) - \beta (x - \mu_x)^2 - \gamma (y - \mu_y)^2 - \delta (x - \mu_x)(y - \mu_y) = 0 \quad (2)$$

となり、上式より $p(\cdot)$ 最大エントロピー推定値

$$p(x, y) = \exp \{ -\alpha - \beta (x - \mu_x)^2 - \gamma (y - \mu_y)^2 - \delta (x - \mu_x)(y - \mu_y) \} \quad (3)$$

が得られる。

ここで、

$$A = \exp(-\alpha) \quad (4)$$

とおき、式(3)を式(3.12)に代入すると次式が成立する。

$$A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \{ -\{ \beta (x - \mu_x)^2 + \gamma (y - \mu_y)^2 + \delta (x - \mu_x)(y - \mu_y) \} \} dx dy = 1 \quad (5)$$

いま、変数変換

$$s = \beta^{1/2} (x - \mu_x), \quad t = \gamma^{1/2} (y - \mu_y) \quad (6)$$

を行うと式(5)の2重積分は

$$\frac{1}{(\beta \gamma)^{1/2}} \int_{-\infty}^{\infty} \exp(-t^2) \int_{-\infty}^{\infty} \exp(-s^2 - \frac{\delta}{(\beta \gamma)^{1/2}} ts) ds dt \quad (7)$$

と変形される。ここで、 $\beta > 0$ ， $\gamma > 0$ とする。

一般に

$$\int_{-\infty}^{\infty} \exp(-p^2 z^2 \pm qz) dz = \exp\left(\frac{q^2}{4p^2}\right) \frac{\pi^{1/2}}{p}, \quad p > 0 \quad (8)^{12)}$$

であるから式(7)の s, t に関する積分を実行すると、式(5) から

$$A = \frac{(\beta \gamma - \delta^2/4)^{1/2}}{\pi} \quad (9)$$

が得られる。ここに、 $\beta \gamma - \delta^2/4 > 0$ とし、この条件は以下の議論でも成立しているものとする。

次に、式(3)を式(3.13)に代入すると次式が成立する。

$$A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)^2 \exp[-\{\beta(x - \mu_x)^2 + \gamma(y - \mu_y)^2 + \delta(x - \mu_x)(y - \mu_y)\}] = \sigma_x^2 \quad (10)$$

ここでも、式(6)の変数変換を行うと上式の2重積分は

$$\frac{1}{\beta^{3/2} \gamma^{1/2}} \int_{-\infty}^{\infty} s^2 \exp(-s^2) \int_{-\infty}^{\infty} \exp(-t^2 - \frac{\delta}{(\beta \gamma)^{1/2}} st) ds dt \quad (11)$$

と変形される。そこで、今度は t に関する積分は式(8)を使って、その後の s に関する積分は

$$\int_{-\infty}^{\infty} z^2 \exp(-pz^2 + 2qz) dz = \frac{1}{2p} (\pi/p)^{1/2} (1 + 2\frac{q^2}{p}) \exp(\frac{q^2}{p}), \quad p > 0 \quad (12)^{13)}$$

を使って実行すると、式(10)は

$$\frac{A \pi}{2 \gamma^{1/2}} \left(\beta - \frac{\delta^2}{4 \gamma} \right)^{-3/2} = \sigma_x^2 \quad (13)$$

となる。

同様に、式(3) と式(3.14)から、式(13)に対応して

$$\frac{A \pi}{2 \beta^{1/2}} \left(\beta - \frac{\delta^2}{4 \beta} \right)^{-3/2} = \sigma_y^2 \quad (14)$$

が求まる。

最後に、式(3) を式(3.15)に代入すると

$$\begin{aligned} & A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) \exp[-\{\beta (x - \mu_x)^2 + \gamma (y - \mu_y)^2 + \delta (x - \mu_x)(y - \mu_y)\}] \\ & = \rho \sigma_x \sigma_y \end{aligned} \quad (15)$$

となるため、やはり式(6) の変数変換を行って上式の2重積分を

$$\frac{1}{\beta \gamma} \int_{-\infty}^{\infty} t \cdot \exp(-t^2) \int_{-\infty}^{\infty} s \cdot \exp(-s^2 - \frac{\delta}{(\beta \gamma)^{1/2}} ts) ds dt \quad (16)$$

と変形する。したがって、ここではsに関する積分は

$$\int_{-\infty}^{\infty} z \cdot \exp(-pz^2 + 2qz) dz = \frac{q}{p} (\pi/p)^{1/2} \exp(\frac{q^2}{p}) \quad , \quad p > 0 \quad (17)^{13)}$$

を使って、その後のtに関する積分は式(12)を使って実行すると式(15)は

$$- \frac{A \pi \delta}{4 (\beta \gamma)^{3/2}} \left(1 - \frac{\delta^2}{4 \beta \gamma} \right)^{-3/2} = \rho \sigma_x \sigma_y \quad (18)$$

となる。

以上の結果から、式(9),(13),(14),(18)を連立させて解くと

$$\begin{aligned}
 A &= \frac{1}{2\pi \sigma_x \sigma_y (1-\rho^2)^{1/2}}, & \beta &= \frac{1}{2\sigma_x^2(1-\rho^2)}, \\
 \gamma &= \frac{1}{2\sigma_y^2(1-\rho^2)}, & \delta &= \frac{-\rho}{\sigma_x \sigma_y (1-\rho^2)}
 \end{aligned}
 \tag{19}$$

が得られ，上式を式(3) に代入すれば，式(3.16)の2変数正規分布が導出される。

Appendix B 2変数最大エントロピー分布からの2変数指数分布（相関が零の場合）の導出

$x[0, \infty)$, $y[0, \infty)$ とし，式(3.19)～(3.21)を制約条件として式(3.1) を最大にする $p(\cdot)$ は，式(3.22)の2変数指数分布になることを示す。

この問題のラグランジュ関数は

$$\begin{aligned}
 L &= H(x, y) + (\alpha - 1) \left\{ 1 - \int_0^\infty \int_0^\infty p(x, y) dx dy \right\} + \\
 &\quad \beta \left\{ \eta_x - \int_0^\infty \int_0^\infty x p(x, y) dx dy \right\} + \gamma \left\{ \eta_y - \int_0^\infty \int_0^\infty y p(x, y) dx dy \right\}
 \end{aligned}
 \tag{1}$$

となる。上式を $p(\cdot)$ について変分をとり零とおくと，

$$-\{1 + \ln p(x, y)\} - (\alpha - 1) - \beta x - \gamma y = 0
 \tag{2}$$

となり，上式より $p(\cdot)$ の最大エントロピー推定値

$$p(x, y) = \exp(-\alpha - \beta x - \gamma y)
 \tag{3}$$

が得られる。

ここで，

$$A = \exp(-\alpha) \quad (4)$$

とおき、式(3) を式(3.19)に代入すると次式が得られる。

$$A \int_0^{\infty} \int_0^{\infty} \exp(-\beta x - \gamma y) dx dy = 1 \quad (5)$$

上式の2重積分を実行して整理すると

$$A = \beta \gamma, \quad \beta > 0, \quad \gamma > 0 \quad (6)$$

となる。

次に、式(3) を式(3.20)に代入すると次式が得られる。

$$A \int_0^{\infty} \int_0^{\infty} x \cdot \exp(-\beta x - \gamma y) dx dy = \eta_x \quad (7)$$

一般に

$$\int_0^{\infty} z^{\alpha} \exp(-az) dz = \Gamma(\alpha+1)/a^{\alpha+1}, \quad \alpha > -1, \quad a > 0 \quad (8)^{14)}$$

であるから、上式を使って式(7) の積分を実行すると

$$A/(\gamma \beta^2) = \eta_x, \quad \beta > 0, \quad \gamma > 0 \quad (9)$$

が得られる。

同様に、式(3) を式(3.21)に代入して積分を実行すると

$$A/(\beta \gamma^2) = \mu_y, \quad \beta > 0, \quad \gamma > 0 \quad (10)$$

が得られる。

そこで、式(6),(9),(10)を連立させて解くと

$$A=1/(\eta_x \eta_y) , \quad \beta=1/\eta_x , \quad \gamma=1/\eta_y \quad (11)$$

となり、上式を式(3)に代入すると、式(3.22)の相関が零の場合の2変数指数分布になる。

参 考 文 献

- 1) 井沢竜夫：二変数の Γ -分布について（降水量分布 第2報），気象と統計，第4巻，第1号，pp.9～15，1953年。
- 2) 井沢竜夫：二変数 Γ -分布について（続），気象と統計，第4巻，第2号，pp.15～19，1953年。
- 3) Freund, J. E. : A bivariate extension of the exponential distribution, American Statistical Association Journal, Vol.56, pp.971～977, Dec., 1961.
- 4) 例えば，長尾正志：水文統計解析，河川工学百年の歩みと淀川，p.88，1978年9月。
- 5) 石原安雄・長尾正志：計画高水流量の二変数確率論的研究，京都大学防災研究所年報，第6号，pp.181～188，1963年7月。
- 6) 長尾正志・角屋 睦：二変数ガンマ分布とその適用に関する研究（1）——二変数指数分布の基礎理論——，京都大学防災研究所年報，第13号，B，pp.105～133，1970年3月。
- 7) 長尾正志・角屋 睦：二変数ガンマ分布とその適用に関する研究（2）——二変数ガンマ分布の母数推定法——，京都大学防災研究所年報，第14号，B，pp.43～56，1971年4月。
- 8) 端野道夫・杉 雄司：複数の二変数指数型分布の結合とその応用に関する研究，徳島大学工学部研究報告，第29号，pp.49～57，1984年。
- 9) 寒川典昭・荒木正夫・寺島 彰：2変数MEP分布とその特性に関する研究，第28回水理講演会論文集，pp.397-402，1984年2月。
- 10) 寒川典昭・荒木正夫・寺島 彰：2変数最大エントロピー分布の適用性，信州大学工学部紀要，第62号，pp.33～48，1987年9月。
- 11) 例えば，笠原芳郎：情報理論と通信方式，共立出版，p.151，1965年10月。
- 12) Gradshteyn, I. S. and Ryzhik, I. M. : Table of integrals, series, and products, Academic Press, Corrected and Enlarged Edition, p.307, 1980.
- 13) 前出 12), p.338.
- 14) 森口繁一，宇田川 久，一松 信：数学公式I——微分積分・平面曲線——，岩波全書，p.230，1956年9月。

第4章 多変数最大エントロピー分布

第1節 概 説

本章は、前章の2変数最大エントロピー分布を3変数以上の多変数の場合に拡張したものである。当然、2変数分布も複数の確率変数を同時に取り扱っているという点で、この分布に含まれるが、従来の水文統計における多変数分布を用いる研究は、主として2変数を中心に展開してきた経緯があるため、前章では特に2変数の場合を取り上げ詳細に議論した。したがって、ここでは、さらに一般化して、任意個数の確率変数を同時に取り扱った多変数最大エントロピー分布について考察する。

一般に、相関のある複数の水文量を同時に考慮して、治水・利水計画を策定すべき場合が多々あり、このとき多変数の同時確率分布が必要となる。例えば長沢¹⁾が述べている、洪水の継続時間、その前後に降る雨（内水降雨）、およびピーク水位による内水排除計画の場合、また、長尾²⁾が指摘しているように、洪水時の、ピーク水位、流量、流速等が破堤・溢流に関係する場合等では、それらの諸量を確率変数とみなした同時確率分布が要求される。石原・長尾³⁾が2変数問題として取り扱った合流河川における治水問題も、合流時差の影響が大きければ、3変数問題として取り扱わなければならない課題である。

ところが、理論面に目を向けると、正規分布以外の多変数確率分布の解明は、ほとんどなされていないのが現状である。数学的な興味においては、Krishnamoorthy・Parthasarathy⁴⁾の多変数Gamma分布の研究があるが、水文統計に導入されるには解決されるべき問題が少なくないようである。しかも、第3章第1節で記述したように、水文量は原型のままで正規分布に従うのは、特殊な場合に限られている。

そこで、本章では、2変数のところで記述したのと同じように、水文量はしばしば小標本であるために有効で、しかも密度関数式が簡潔で利用しやすい、多変数最大エントロピー分布を提案する⁵⁾。第2節では、2変数の場合を拡張した多変数最大エントロピー分布を定式化し、その分布からの既存の多変数分布の分布の導出について触れるとともに、パラメタ同定法を示す。第3節では、仮定した母集団に対する、真のモーメントを使った場合の最大エントロピー分布の適用性を検討する。ただし、シミュレーションデータに対する推定分布の適用性と、モーメントの変動に伴う推定分布の感度分析は、母集団として正規分布しか利用できないこと、およびその検討結果が2変数の場合から類推できること、の理由により省略している。最後に、第5節で、上述の石原・長尾の取り扱った問題と同様に、3地点流量データを用いた3変

数最大エントロピー分布と，渇水問題に対処するための3地点8月降水データによる3変数最大エントロピー分布を求め，水文量に対する適用性について確かめる。

第2節 確率密度関数の推定

2.1 定式化

第3章2.1で議論した2変数での定式化を， $n(\geq 3)$ 変数に拡張することは容易である。ここでは，その概略を説明する。

いま， n 個の確率変数を x_1, x_2, \dots, x_n と書き，その同時確率密度関数を $p(x_1, x_2, \dots, x_n)$ とすると，多変数で定義されたエントロピーは

$$H(x_1, x_2, \dots, x_n) = - \int \cdots \int p(x_1, x_2, \dots, x_n) \ln p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \quad (4.1)$$

となる。⁶⁾ 一方，式(3.2)，式(3.3)に対応した，確率密度関数が具備すべき条件と情報の与え方の表現式は，次の1式で書かれる。

$$\begin{aligned} \int \cdots \int g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \\ = E[g_r(x_1, x_2, \dots, x_n)], \quad r=1, 2, \dots, N \end{aligned} \quad (4.2)$$

当然，上式は確率密度関数が具備すべき条件 $g_r(x_1, x_2, \dots, x_n)=1$ を含む。第2章2.1と第3章2.1ではこの条件を別個に取り出したのに対し，ここではそれを式(4.2)に含めたのは，本章2.3の $g_r(x_1, x_2, \dots, x_n) = x_1^{b_1} x_2^{b_2}, \dots, x_n^{b_n}$ の中にこの条件を含めた方が，相関構造の記述が簡明になるからである。

そこで，第3章2.1で記述した立場から，ここでは式(4.2)を制約条件として式(4.1)を最大にする $p(\cdot)$ をラグランジュの未定乗数法で求めると，式(3.4)～(3.6)に対応した，ラグランジュ関数の式，その変分を零と置いた式，および最大エントロピー推定値として，次の3式が得られる。

$$L = H(x_1, x_2, \dots, x_n) + \sum_{r=1}^N \lambda_r \{E[g_r(x_1, x_2, \dots, x_n)] - \int \dots \int g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n\} \quad (4.3)$$

$$-\{1 + \ln p(x_1, x_2, \dots, x_n)\} - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n) = 0 \quad (4.4)$$

$$p(x_1, x_2, \dots, x_n) = \exp\{-1 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\} \quad (4.5)$$

式(4.5)は、一般的な多変数最大エントロピー分布である。もちろん、ラグランジュ乗数 λ_r は推定分布のパラメタであること、その数は $(N-1)$ 個であること、および $g_r(\cdot)$ として種々の関数を与えることができることは、1変数、2変数の場合と同様である。

2.2 分布の導出

多変数最大エントロピー分布も1変数、2変数の場合と同じように形式的には無限個の分布形を記述することができる。ここでは、第2章、第3章で、最大エントロピー分布を水文統計に導入するときの思想的な背景となった一様分布について、制約条件と得られた分布形を多変数の場合で示すと以下ようになる。

$$g_1(x_1, x_2, \dots, x_n) = 1 \quad (4.6)$$

とし、確率変数の存在範囲を $x_1 [1L_1, 1L_2]$, $x_2 [2L_1, 2L_2]$, \dots , $x_n [nL_1, nL_2]$ として式(4.2)を次のように書き換える。

$$\int_{nL_1}^{nL_2} \dots \int_{2L_1}^{2L_2} \int_{1L_1}^{1L_2} p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1, \\ 1L_2 > 1L_1, 2L_2 > 2L_1, \dots, nL_2 > nL_1 \quad (4.7)$$

このとき、式(4.5)は

$$p(x_1, x_2, \dots, x_n) = \frac{1}{(l_2 - l_1)(l_2 - l_1) \cdots (l_2 - l_1)} \quad (4.8)$$

となり、これは多変数一様分布である。

また、相関が零の場合の多変数指数分布も簡単に導出できるが、前章2.2 3)と同様であるので、省略する。一方、多変数正規分布の場合、前章のAppendix Aで記述したような、ラグランジュ乗数の決定が難解であるため、今のところ導出されていない。

なお、ここでは既存分布の導出という立場にあるので、少数個の分布を対象としているが、水文統計において上述の分布の適用場面は極めて限られる。ところが、最大エントロピー分布は、式(4.5)の $g_r(\cdot)$ に具体的な関数形を与えることによって、既存分布にない多くの分布形を見出すことが可能である。したがって、水文統計は、こうして得られた分布を用いて議論されることが望まれ、これが実用面においても最大エントロピー分布に高い評価を与える一因と考えられる。

2.3 パラメタ同定法

ここでも、第2章 2.4、および第3章2.3で記述した理由により、統計モーメントで情報を与えることを考える。 n 個の確率変数の存在範囲を $[0, \infty)$ とすると、式(4.2)は次のように書き換えられる。

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_1^{a_1} p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = \mu_{a_1}, \quad a_1 = 1, 2, \dots, Na_1 \quad (4.9.1)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_2^{a_2} p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = \mu_{a_2}, \quad a_2 = 1, 2, \dots, Na_2 \quad (4.9.2)$$

...

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_n^{a_n} p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = \mu_{a_n}, \quad a_n = 1, 2, \dots, Na_n \quad (4.9.n)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} p(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n = \mu_{b_1 b_2 \cdots b_n},$$

$$b_1=0, 1, \cdots, Nb_1, \quad b_2=0, 1, \cdots, Nb_2, \quad \cdots, \quad b_n=0, 1, \cdots, Nb_n \quad (4.9.n+1)$$

ここに、 μ_{a_i} は x_i に関する原点のまわりの a_i 次モーメント、 $\mu_{b_1 b_2 \cdots b_n}$ は x_1, x_2, \cdots, x_n に関する原点のまわりの b_1 次、 b_2 次、 \cdots 、 b_n 次の相乗モーメントである。 $b_1=b_2=\cdots=b_n=0$ は確率密度関数が具備すべき条件であり、 $\mu_{00\cdots0}=1$ とする。また b_i の 1 つが 1 以上の整数で、他は零となることはない。なぜなら、このような情報は式 (4.9.1) ~ (4.9.n) に含まれているからである。式 (4.9.1) ~ (4.9.n+1) を制約条件とした多変数最大エントロピー分布は、式 (4.5) より次式となる。

$$p(x_1, x_2, \cdots, x_n) = \exp\left(-\sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n}\right)$$

$$-\sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \quad (4.10)$$

ここに、 γ_{a_i} 、 $\delta_{b_1 b_2 \cdots b_n}$ は分布のパラメタであり、式 (4.5) の λ_r に相当するが、2 変数の場合と同様、確率変数との関係を明示するためにパラメタの表示を変更した。

上式を式 (4.9.1) ~ (4.9.n+1) に代入すると次式が得られる。

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_1^{a_1} \exp\left(-\sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n}\right)$$

$$-\sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} dx_1 dx_2 \cdots dx_n = \mu_{a_1},$$

$$a_1=1, 2, \cdots, Na_1 \quad (4.11.1)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_2^{a_2} \exp\left(-\sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n}\right)$$

$$-\sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} dx_1 dx_2 \cdots dx_n = \mu_{a_2},$$

$$a_2=1, 2, \cdots, Na_2 \quad (4.11.2)$$

...

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_n^{a_n} \exp\left(-1 - \sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n}\right) \\ - \sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}) dx_1 dx_2 \cdots dx_n = \mu_{a_n}, \\ a_n = 1, 2, \cdots, Na_n \quad (4.11.n)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \exp\left(-1 - \sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n}\right) \\ - \sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}) dx_1 dx_2 \cdots dx_n \\ = \mu_{b_1 b_2 \cdots b_n}, \quad b_1 = 0, 1, \cdots, Nb_1, \quad b_2 = 0, 1, \cdots, Nb_2, \cdots, \quad b_n = 0, 1, \cdots, Nb_n \quad (4.11.n+1)$$

したがって、式(4.10)のパラメタ γ_{a_i} , $\delta_{b_1 b_2 \cdots b_n}$ は式(4.11.1)~(4.11.n+1)を連立させて解くことにより決定される。

ここで、式(4.11.1)~(4.11.n+1)は非線形であるため、 γ_{a_i} , $\delta_{b_1 b_2 \cdots b_n}$ に関して解くために第3章2.3 で用いた方法を多変数の場合に拡張する。

いま、 γ_{a_i} , $\delta_{b_1 b_2 \cdots b_n}$ の近似値として $\gamma_{a_i}^0$, $\delta_{b_1 b_2 \cdots b_n}^0$ を与え、残差を ϵ_{a_i} , $\epsilon_{b_1 b_2 \cdots b_n}$ とおくと次式が成立する。

$$\gamma_{a_1} = \gamma_{a_1}^0 + \epsilon_{a_1}, \quad \gamma_{a_2} = \gamma_{a_2}^0 + \epsilon_{a_2}, \quad \cdots, \quad \gamma_{a_n} = \gamma_{a_n}^0 + \epsilon_{a_n}, \\ \delta_{b_1 b_2 \cdots b_n} = \delta_{b_1 b_2 \cdots b_n}^0 + \epsilon_{b_1 b_2 \cdots b_n} \quad (4.12)$$

そこで、上式を式(4.11.1)~(4.11.n+1)に代入して、 ϵ_{a_i} , $\epsilon_{b_1 b_2 \cdots b_n}$ に関して Taylor 展開し、 $(\epsilon_{a_i})^2$, $(\epsilon_{b_1 b_2 \cdots b_n})^2$ 以上の微小量の項を無視して整理すると次式が得られる。

$$\sum_{j_1=1}^{Na_1} A^{i_1+j_1}_{i_1} \in j_1 + \sum_{i_2=1}^{Na_2} A^{i_1}_{i_1} j_2 \in j_2 + \cdots + \sum_{i_n=1}^{Na_n} A^{j_2}_{i_n} j_n \in i_n$$

$$+ \sum_{i_n=0}^{Nb_n} \cdots \sum_{i_2=0}^{Nb_2} \sum_{j_1=0}^{Nb_1} A^{i_1+j_1}_{i_1} j_2 \cdots i_n \in j_1 i_2 \cdots i_n = B^{i_1}_{i_1}, \quad i_1=1, 2, \cdots, Na_1 \quad (4.13.1)$$

$$\sum_{i_1=1}^{Na_1} A^{j_1}_{i_1} j_2 \in i_1 + \sum_{j_2=1}^{Na_2} A^{i_2+j_2}_{i_2} \in j_2 + \cdots + \sum_{i_n=1}^{Na_n} A^{j_2}_{i_n} j_n \in i_n$$

$$+ \sum_{i_n=0}^{Nb_n} \cdots \sum_{j_2=0}^{Nb_2} \sum_{i_1=0}^{Na_1} A^{j_1}_{i_1} i_2+j_2 \cdots i_n \in i_1 j_2 \cdots i_n = B^{j_2}_{i_2}, \quad i_2=1, 2, \cdots, Na_2 \quad (4.13.2)$$

...

$$\sum_{i_1=1}^{Na_1} A^{j_1}_{i_1} i_n \in i_1 + \sum_{i_2=1}^{Na_2} A^{j_2}_{i_2} i_n \in j_2 + \cdots + \sum_{j_n=1}^{Na_n} A^{i_n+j_n}_{i_n} \in j_n$$

$$+ \sum_{j_n=0}^{Nb_n} \cdots \sum_{i_2=0}^{Nb_2} \sum_{i_1=0}^{Na_1} A^{j_1}_{i_1} j_2 \cdots i_n+j_n \in i_1 i_2 \cdots j_n = B^{i_n}_{i_n}, \quad i_n=1, 2, \cdots, Na_n \quad (4.13.n)$$

$$\sum_{j_1=1}^{Na_1} A^{i_1+j_1}_{i_1} j_2 \cdots i_n \in j_1 + \sum_{j_2=1}^{Na_2} A^{i_1}_{i_1} i_2+j_2 \cdots j_n \in j_2 + \cdots$$

$$+ \sum_{j_n=1}^{Na_n} A^{i_1}_{i_1} j_2 \cdots i_n+j_n \in j_n + \sum_{j_n=0}^{Nb_n} \cdots \sum_{j_2=0}^{Nb_2} \sum_{j_1=0}^{Nb_1} A^{i_1+j_1}_{i_1} i_2+j_2 \cdots i_n+j_n$$

$$\in j_1 j_2 \cdots j_n = B^{i_1}_{i_1} i_2 \cdots i_n,$$

$$i_1=0, 1, \cdots, Nb_1, \quad i_2=0, 1, \cdots, Nb_2, \quad \dots, \quad i_n=0, 1, \cdots, Nb_n \quad (4.13.n+1)$$

ここに,

$$A_k^{i_k} = \int_0^\infty \cdots \int_0^\infty \int_0^\infty x_k^{i_k} \exp\left(-1 - \sum_{a_1=1}^{Na_1} \gamma_{a_1}^0 x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2}^0 x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n}^0 x_n^{a_n}\right) \\ - \sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1}^0 b_2 \cdots b_n x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} dx_1 dx_2 \cdots dx_n \\ k=1, 2, \cdots, i_k=0, 1, \cdots \quad (4.14)$$

$$A_1^{i_1} 2^{i_2} \cdots n^{i_n} = \int_0^\infty \cdots \int_0^\infty \int_0^\infty x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} \exp\left(-1 - \sum_{a_1=1}^{Na_1} \gamma_{a_1}^0 x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2}^0 x_2^{a_2} \right. \\ \left. - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n}^0 x_n^{a_n} - \sum_{b_n=0}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1}^0 b_2 \cdots b_n x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}\right) dx_1 dx_2 \cdots dx_n, \\ i_1=0, 1, \cdots, i_2=0, 1, \cdots, \cdots, i_n=0, 1, \cdots \quad (4.15)$$

である。ただし、式 (4.13.1) ~ (4.13.n) で $A_s^{i_s} t^{i_t}$ は $A_1^{i_1} \cdots s^{i_s} s^{i_s} s^{i_s+1} \cdots t^{i_t} t^{i_t+1} \cdots n^{i_n}$ ($i_s, i_t \geq 1$) の略記である。また

$$B_s^{i_s} = A_s^{i_s} - s \mu_{i_s}, \quad B_1^{i_1} 2^{i_2} \cdots n^{i_n} = A_1^{i_1} 2^{i_2} \cdots n^{i_n} - \mu_{i_1 i_2 \cdots i_n} \quad (4.16)$$

である。

式 (4.13.1) ~ (4.13.n+1) は $i \in a_i, \epsilon \in b_1 b_2 \cdots b_n$ に関する連立 1 次方程式であり、元の数は式 (4.9.n+1) の相関構造の取り入れ方に依存して決まるが、どのような取り入れ方を採用するにせよ変数の数と方程式の数が一致するので解き得る。

実際、式 (4.10) のパラメタは、式 (4.13.1) ~ (4.13.n+1) を繰り返し解くことにより求められる。具体的なパラメタ同定手順は、収束判定基準についての考察も含めて、前章 2.3 とほとんど同じであるので、その記述を省略する。ただし、ここでは、p.52 の 6) に関する計算は必要としない。

2.4 モーメントの基準化

ここでも、第 2 章 2.6 で記述した理由により、統計モーメントの基準化を考える。

確率変数 u_1, u_2, \cdots, u_n の平均が 1 となるような変数変換、すなわち

$$u_1 = x_1 / ({}_1\mu_1), \quad u_2 = x_2 / ({}_2\mu_1), \quad \dots, \quad u_n = x_n / ({}_n\mu_1) \quad (4.17)$$

を行うと, u_1, u_2, \dots, u_n に関する確率密度関数 $q(u_1, u_2, \dots, u_n)$ は, Jacobian $J(\cdot)$ を介して

$$q(u_1, u_2, \dots, u_n) = J\left(\frac{x_1, x_2, \dots, x_n}{u_1, u_2, \dots, u_n}\right) p(x_1, x_2, \dots, x_n) \quad (4.18)$$

となる。ここに,

$$J\left(\frac{x_1, x_2, \dots, x_n}{u_1, u_2, \dots, u_n}\right) = \begin{vmatrix} {}_1\mu_1 & 0 & \dots & 0 \\ 0 & {}_2\mu_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & {}_n\mu_1 \end{vmatrix} = {}_1\mu_1 \cdot {}_2\mu_1 \cdot \dots \cdot {}_n\mu_1 \quad (4.19)$$

であるから, 式 (4.18) は

$$q(u_1, u_2, \dots, u_n) = p(x_1, x_2, \dots, x_n) \cdot {}_1\mu_1 \cdot {}_2\mu_1 \cdot \dots \cdot {}_n\mu_1 \quad (4.20)$$

あるいは,

$$p(x_1, x_2, \dots, x_n) = q(u_1, u_2, \dots, u_n) / ({}_1\mu_1 \cdot {}_2\mu_1 \cdot \dots \cdot {}_n\mu_1) \quad (4.21)$$

と表現できる。

上述の変数変換を行うと, 式 (4.9.1) ~ (4.9.n+1) は次のように書き換えられる。

$$\begin{aligned} & \int_0^\infty \dots \int_0^\infty \int_0^\infty u_1^{a_1} q(u_1, u_2, \dots, u_n) du_1 du_2 \dots du_n \\ &= {}_1\mu_{a_1} / ({}_1\mu_1)^{a_1} \equiv {}_1m_{a_1}, \quad a_1 = 1, 2, \dots, Na_1 \end{aligned} \quad (4.22.1)$$

$$\begin{aligned} & \int_0^\infty \dots \int_0^\infty \int_0^\infty u_2^{a_2} q(u_1, u_2, \dots, u_n) du_1 du_2 \dots du_n \\ &= {}_2\mu_{a_2} / ({}_2\mu_1)^{a_2} \equiv {}_2m_{a_2}, \quad a_2 = 1, 2, \dots, Na_2 \end{aligned} \quad (4.22.2)$$

...

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty u_n^{a_n} q(u_1, u_2, \dots, u_n) du_1 du_2 \cdots du_n$$

$$= {}_n \mu_{a_n} / ({}_n \mu_1)^{a_n} \equiv {}_n m_{a_n}, \quad a_n = 1, 2, \dots, N_{a_n} \quad (4.22.n)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty u_1^{b_1} u_2^{b_2} \cdots u_n^{b_n} q(u_1, u_2, \dots, u_n) du_1 du_2 \cdots du_n$$

$$= \mu_{b_1} b_2 \cdots b_n / \{({}_1 \mu_1)^{b_1} ({}_2 \mu_1)^{b_2} \cdots ({}_n \mu_1)^{b_n}\} \equiv m_{b_1 b_2 \cdots b_n},$$

$$b_1 = 0, 1, \dots, N_{b_1}, \quad b_2 = 0, 1, \dots, N_{b_2}, \quad \dots, \quad b_n = 0, 1, \dots, N_{b_n} \quad (4.22.n+1)$$

ここに，2変数の場合と同様，式(4.22.1)～(4.22.n)で定義されたモーメントを，それぞれ， x_1, x_2, \dots, x_n に関する基準化されたモーメント，式(4.22.n+1)で定義されたモーメントを， x_1, x_2, \dots, x_n に関する基準化された相乗モーメントと呼ぶことにする。

次に， $p(\cdot)$ と $q(\cdot)$ のエントロピーの関係が

$$H(x_1, x_2, \dots, x_n) = H(u_1, u_2, \dots, u_n) + \ln({}_1 \mu_1 \cdots {}_2 \mu_1 \cdots {}_n \mu_1) \quad (4.23)$$

になることは，式(2.49)，(3.46)と同様であり，ここでも実際の適用では式(4.22.1)～(4.22.n+1)を制約条件として $H(u_1, u_2, \dots, u_n)$ を最大にする $q(\cdot)$ を求めておき，しかる後に式(4.21)により $p(\cdot)$ に変換すればよいことは1変数，2変数のところで述べたとおりである。

第3節 既存分布を母集団とした場合の適用性の検討

3.1 仮定した母集団

ここでは，3変数以上の多変数を母集団に仮定する必要がある。仮定母集団としては，前章3.1.1で記述したように，非対称形を含んだ多くの形状を取り得ることが望ましい。しかし，多変数分布が，第1節で述べた現状にある以上，非対称形を示す母集団を用いることは難しいため，本節では多変数正規分布を母集団として採用する。

一般に，多変数正規分布の密度関数は次式で与えられる。⁷⁾

$$P(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \sigma^{ij} (x_i - \mu_i)(x_j - \mu_j) \right\} \quad (4.24)$$

ここに,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix} \quad (4.25)$$

は分散・共分散行列, σ^{ij} は Σ の逆行列の元, μ_i は x_i の平均である。

一方, この分布のモーメント母関数は

$$\begin{aligned} \phi(\theta_1, \theta_2, \dots, \theta_n) &= E[\exp(\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n)] \\ &= \exp\left(\sum_{i=1}^n \mu_i \theta_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} \theta_i \theta_j\right) \end{aligned} \quad (4.26)^8)$$

となるから, これより原点のまわりのモーメントが計算される。

いま, 3変数 x_1, x_2, x_3 の場合, $\mu_1 = \mu_2 = \mu_3 = \mu$, $\sigma_{11} = \sigma_{22} = \sigma_{33} = \sigma$, $\sigma_{12} = \sigma_{23} = \sigma_{31} = \rho\sigma$ とおくと, 3変数正規分布は, 平均 μ , 標準偏差 σ , 相関係数 ρ の3母数で規定されるため, これを $3N(\mu, \sigma, \rho)$ と略記する。また, 一般的な3変数最大エントロピー分布の具体的なモーメントは, 式(4.26)から計算される。

3.2 適用性の検討

ここでは, 3変数最大エントロピー分布が, どの程度母集団を再現できるかを明らかにするため, 前項で定義した $3N(\mu, \sigma, \rho)$ を母集団と仮定する。なお, 最大エントロピー分布の確率変数は $[0, \infty)$ で定義されているので, 母集団とした分布の確率変数の負値領域の影響を小さくするため, μ/σ を大きくすることを考え, 以下の計算では $\mu/\sigma = 2$ として, $2N(2, 1,$

0.2), $2N(2,1,0.5)$, $2N(2,1,0.8)$ の3つの場合を検討した。一方, 3変数最大エントロピー分布は, 式(4.10)より

$$P(x_1, x_2, x_3) = \exp \left(-1 - \sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \sum_{a_3=1}^{Na_3} \gamma_{a_3} x_3^{a_3} \right. \\ \left. - \sum_{b_1=0}^{Nb_1} \sum_{b_2=0}^{Nb_2} \sum_{b_3=0}^{Nb_3} \delta_{b_1 b_2 b_3} x_1^{b_1} x_2^{b_2} x_3^{b_3} \right) \quad (4.27)$$

となり, これを $3M(s, t, u, v)$ と略記する。ここで, $s=Na_1, t=Na_2, u=Na_3$ である。 v は採用した相乗モーメントの指標であり, 本項では以下のようになっている。

1) $v=1$ の場合

採用するモーメント = $\{\mu_{000}=1\}$

2) $v=2$ の場合

採用するモーメント = $\{\mu_{000}, \mu_{110}, \mu_{011}, \mu_{101}\}$

3) $v=3$ の場合

採用するモーメント = $\{\mu_{000}, \mu_{110}, \mu_{011}, \mu_{101}, \mu_{111}\}$

表4.1 は, 上述の母集団の基準化されたモーメントから得られた最大エントロピー分布のパラメタの値である。たとえば, (a) $3M(2,2,2,2)$ の場合, この最大エントロピー分布は次式を表現している。

$$p(u_1, u_2, u_3) = \exp(-1 + 2.0284u_1 - 1.8054u_1^2 + 2.0284u_2 - 1.8054u_2^2 + 2.0284u_3 \\ - 1.8054u_3^2 - 2.573 + 0.7066u_2 u_3 + 0.7066u_1 u_3 + 0.7066u_1 u_2) \quad (4.28)$$

図4.1 は, 母集団と最大エントロピー分布について, 3次元確率密度関数 $p(u_1, u_2, u_3)$ を u_1-u_2 平面の2次元周辺確率密度関数 $p_*(u_1, u_2)$ に変換し, 母集団と $3M(2,2,2,2)$ の3次元立体図を示したものである。図4.2 は, $p_*(u_1, u_2)$ に関する母集団と $3M(2,2,2,2), 3M(3,3,3,2)$ のメッシュ確率である。適合度の改善は, メッシュごとで検討し, 改善されたメッシュの数およびその改善の度合から判断した。このことは, 次節においても同様である。

母集団に対する $3M(2,2,2,2)$ の適合度は原点付近で劣るが, これは母集団分布の確率変数の

表4.1 $3N(\mu, \sigma, \rho)$ を母集団とした場合の最大エントロピー分布のパラメタの値 ($1\gamma. = 2\gamma. = 3\gamma. = 4\gamma.$)

(a) Population : $3N(2, 1, 0.2)$

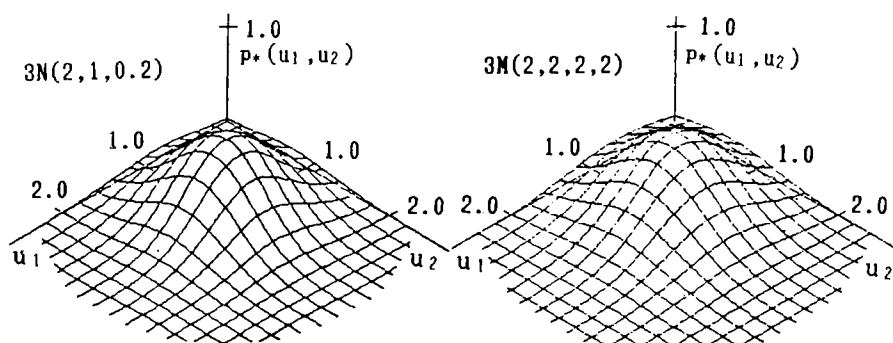
	$3M(2, 2, 2, 2)$	$3M(3, 3, 3, 2)$	$3M(3, 3, 3, 3)$
$1f_1$	-2.0284	1.5591	1.5589
$1f_2$	1.8054	-2.3145	-2.3145
$1f_3$	—	1.2997	1.2997
δ_{000}	2.5723	0.3950	0.3952
δ_{011}	-0.7066	-0.7119	-0.7117
δ_{101}	-0.7066	-0.7119	-0.7117
δ_{110}	-0.7066	-0.7119	-0.7117
δ_{111}	—	—	-0.0002

(b) Population : $3N(2, 1, 0.5)$

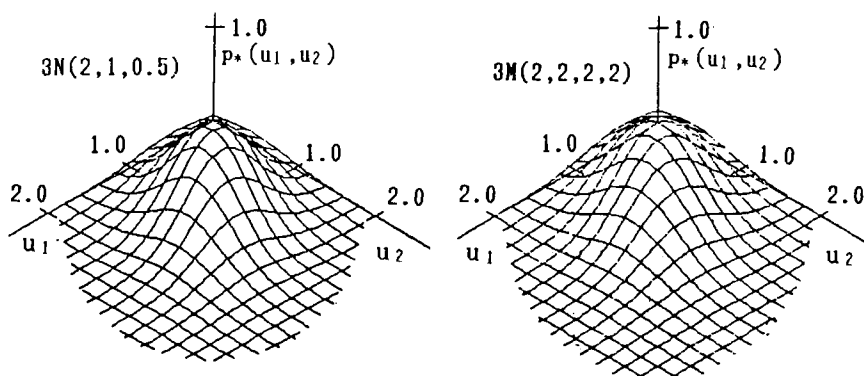
	$3M(2, 2, 2, 2)$	$3M(3, 3, 3, 2)$	$3M(3, 3, 3, 3)$
$1f_1$	-1.2519	2.3625	2.3584
$1f_2$	2.6605	-1.4717	-1.4841
$1f_3$	—	1.3073	1.3091
δ_{000}	1.1332	-1.0948	-1.0905
δ_{011}	-1.9519	-1.9631	-1.9538
δ_{101}	-1.9519	-1.9631	-1.9538
δ_{110}	-1.9519	-1.9631	-1.9538
δ_{111}	—	—	-0.0093

(c) Population : $3N(2, 1, 0.8)$

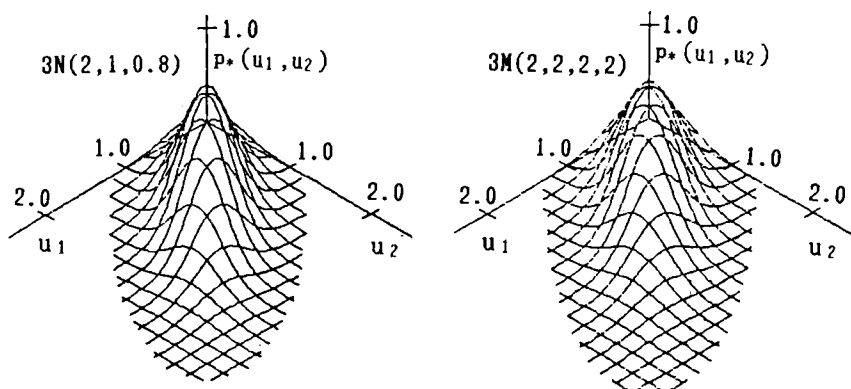
	$3M(2, 2, 2, 2)$	$3M(3, 3, 3, 2)$	$3M(4, 4, 4, 3)$
$1f_1$	-0.9746	2.2588	2.6460
$1f_2$	5.9677	2.1327	2.1424
$1f_3$	—	1.1733	1.3275
$1f_4$	—	—	-0.0648
δ_{000}	0.0138	-1.9805	-2.2246
δ_{011}	-5.4070	-5.2785	-5.8713
δ_{101}	-5.4070	-5.2785	-5.8713
δ_{110}	-5.4070	-5.2785	-5.8713
δ_{111}	—	—	0.5518



(a) Population : $3N(2,1,0.2)$

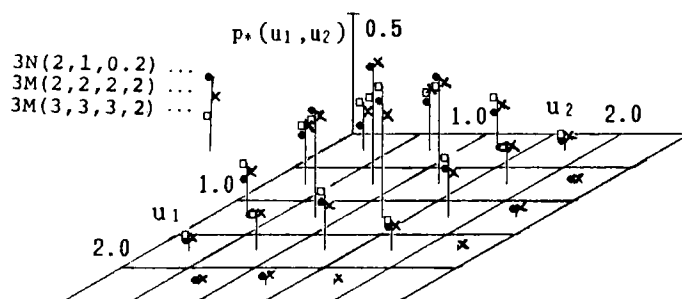


(b) Population : $3N(2,1,0.5)$

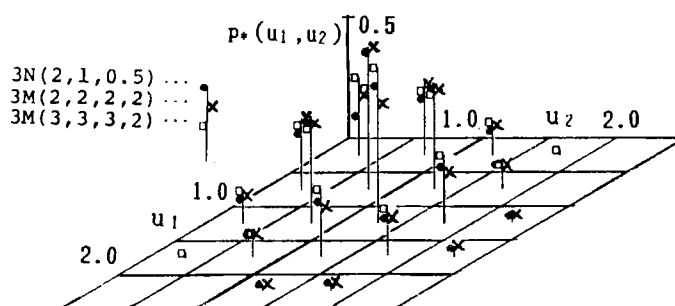


(c) Population : $3N(2,1,0.8)$

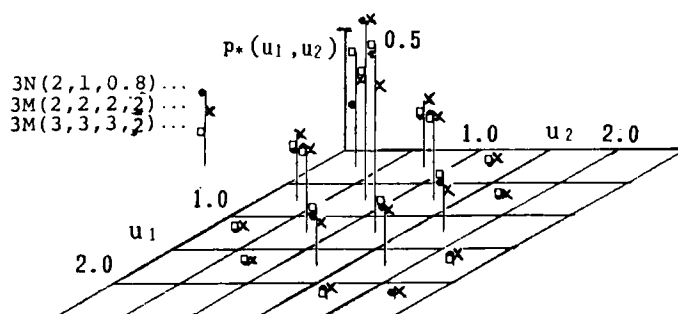
図4.1 $3N(\mu, \sigma, \rho)$ と $3M(2,2,2,2)$ の3次元立体図



(a) Population : $3N(2,1,0.2)$



(b) Population : $3N(2,1,0.5)$



(c) Population : $3N(2,1,0.8)$

図4.2 $3N(\mu, \sigma, \rho)$, $3M(2,2,2,2)$ および $3M(3,3,3,2)$
のメッシュ確率

定義域が $(-\infty, \infty)$ であるのに対し、最大エントロピー分布では $[0, \infty)$ としたためと推察される。しかし、全体的には両者はよく一致しているとともに、確率変数の大きな領域でも、よい合致度を示している。このことは、3変数正規分布が平均と分散・共分散行列によって確定され、そのパラメタは相乗を含めた2次モーメントまでで同定されることに起因している。さらに高次モーメントまで取り入れた適合度検定、すなわち、 $3M(3,3,3,2), 3M(3,3,3,3), 3M(4,4,4,3)$ では、明確な適合度の改善は得られなかった。

以上より、3変数正規分布のような母集団に対しては、相関係数に相当する2次の相乗モーメントを含む、すべての2次以下のモーメントを使った $3M(2,2,2,2)$ で十分よく再現できることがわかった。

第4節 実データへの適用

1変数、2変数での検討を踏まえて、ここでも最大限4次モーメントまでを採用することとした。

4.1 3地点流量の場合

前章4.1で触れた石原・長尾³⁾の合流河川問題は、各支川の流量を2変数にとり、合流後の流量は各支川の流量の線形和で近似して2変数統計に持ち込んだものであるが、本来3変数統計問題である。そこで、ここでは、合流後の流量も確率変数とみなした3変数最大エントロピー分布を求めた。

3地点流量データは、前章4.1で用いたのと同じ、立ヶ花29年間の年最大流量と、それに対応する杭瀬下、小市の最大流量である。それぞれ、基準化されたモーメントに対応する確率変数を u_1, u_2, u_3 とした。表4.2は、基準化されたモーメントから得られた最大エントロピー分布のパラメタの値である。図4.3は1地点で積分した他の2地点の $3M(2,2,2,2)$ の3次元立体図、図4.4は、やはり2変数の周辺分布に変形した、 $3M(2,2,2,1), 3M(2,2,2,2)$ 、およびデータのメッシュ確率である。

杭瀬下-小市、立ヶ花-小市間では、 $3M(2,2,2,1)$ に相乗モーメントの導入やモーメントの次数の増加をはかっても、適合度の改善は明確にならなかった。立ヶ花-杭瀬下間では、 $3M(2,2,2,1)$ から $3M(2,2,2,2)$ にすると適合度が改善され、確率変数がともに大きい領域のデータも説明できるように分布形状の移行がみられるが、さらに $3M(3,3,3,2)$ にしても、その後の改善は見られなかった。また、4次モーメントを制約条件に加えた場合、パラメタ同定がなさ

表4.2 3地点流量データより求めた最大
エントロピー分布のパラメタの値

	3M(2,2,2,1)	3M(2,2,2,2)	3M(3,3,3,2)
$1f_1$	-2.1499	-0.9424	-1.9526
$1f_2$	1.2043	14.1640	15.1870
$1f_3$	—	—	-0.2039
$2f_1$	0.5880	1.9152	3.1241
$2f_2$	0.1179	3.2944	2.4388
$2f_3$	—	—	0.1763
$3f_1$	-4.4673	-3.6575	-3.4942
$3f_2$	2.2809	6.2067	6.0960
$3f_3$	—	—	0.0427
δ_{000}	2.8746	0.0773	0.0283
δ_{011}	—	5.5192	5.5212
δ_{101}	—	-14.1690	-14.2670
δ_{110}	—	-13.1960	-13.3790

れなかった。

以上より、この例では、データ数が少ないため詳細な適合度の議論ができないこと、およびまだ改善の余地が残されていることの問題点はあるが、3M(2,2,2,2) で一応満足される適合度を得られたと判断した。

4.2 3地点8月降水の場合

3地点8月降水として、長野、名古屋、東京の92年間のデータを用い、それぞれの基準化されたモーメントに対応する確率変数を u_1, u_2, u_3 とした。表4.3 は、基準化されたモーメントから得られた最大エントロピー分布のパラメタの値である。図4.5 は図4.3 と同様にして求めた3M(4,4,4,2) の3次元立体図、図4.6 は図4.4 に相当する、3M(2,2,2,2), 3M(4,4,2,2), およびデータのメッシュ確率である。

名古屋－東京間では、3M(2,2,2,1) から3M(2,2,2,2) にすると適合度が改善されたが、3M(3,3,3,2), 3M(4,4,4,2) にしても、その後の改善が明確にならなかった。長野－東京間では、3M(2,2,2,1) から3M(2,2,2,2), 3M(3,3,3,2) にしても適合度の改善は明確にならなかったが、3M(4,4,4,2) にするとやや改善がみられた。長野－名古屋間では、3M(2,2,2,1) から3M(2,2,

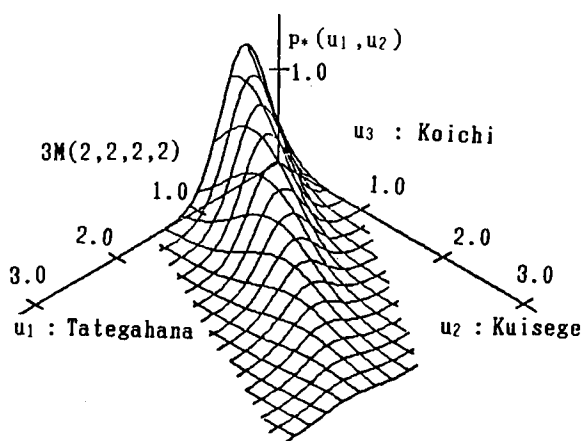
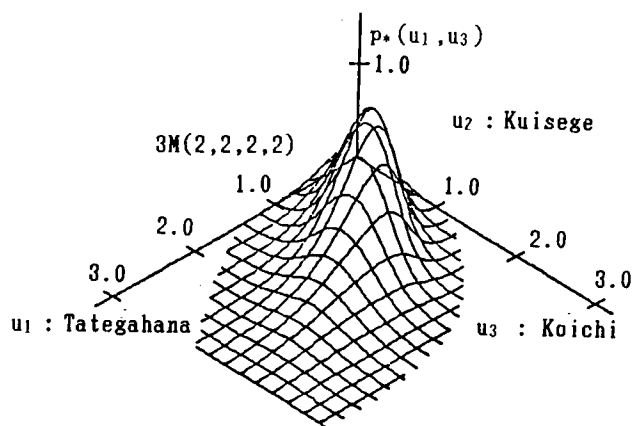
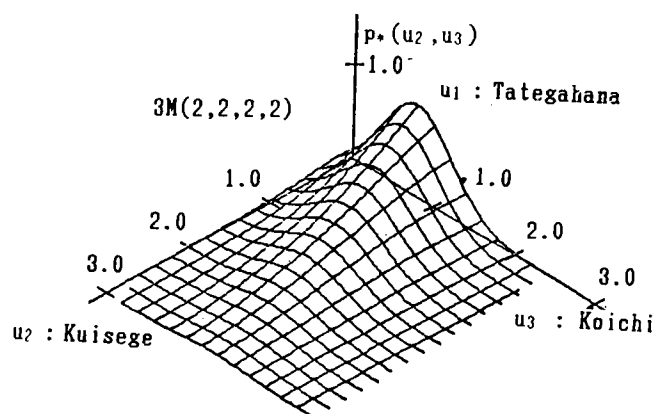


図4.3 3地点流量データより求めた $3M(2,2,2,2)$ の3次元立体図

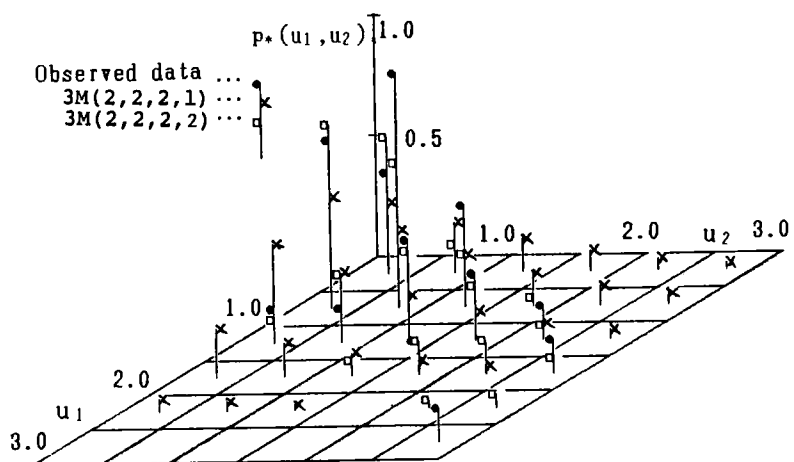
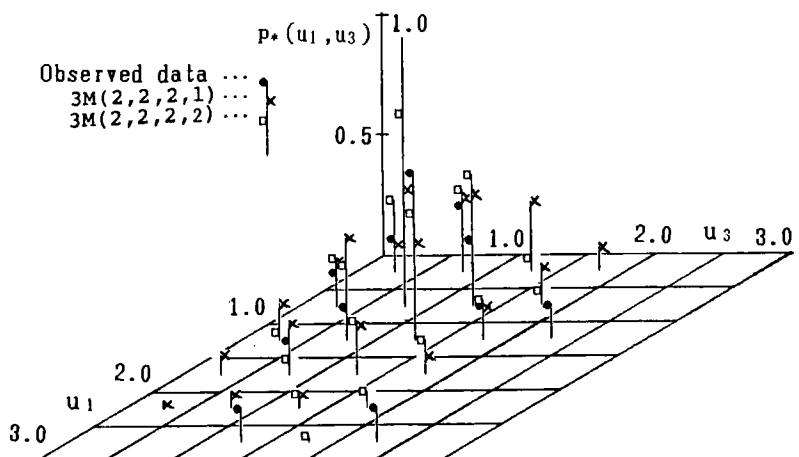
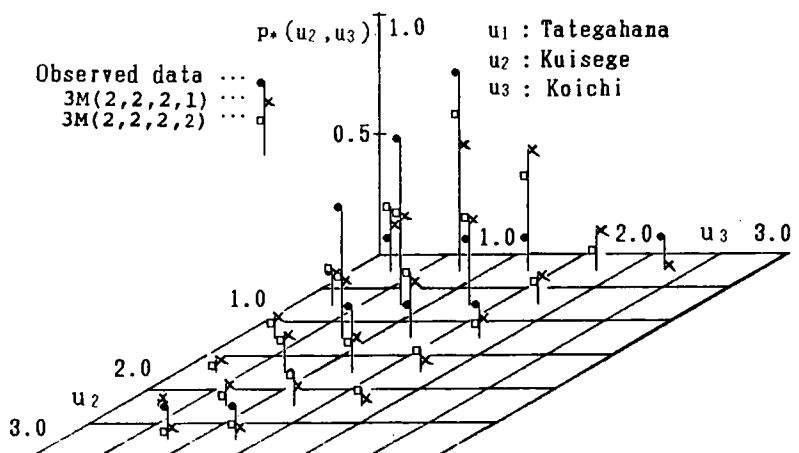


図4.4 $3M(2,2,2,1)$, $3M(2,2,2,2)$ およびデータのメッシュ確率

表4.3 3地点8月降水データより求めた最大
エントロピー分布のパラメタの値

	3M(2,2,2,1)	3M(2,2,2,2)	3M(3,3,3,2)	3M(4,4,4,2)
1f ₁	-3.2701	-2.5037	-2.8257	-9.2945
1f ₂	1.7161	1.9741	2.2756	11.4190
1f ₃	—	—	-0.0827	-4.8645
1f ₄	—	—	—	0.8140
2f ₁	-1.0597	0.2484	0.1661	-2.4341
2f ₂	0.7335	1.1129	1.1807	4.7111
2f ₃	—	—	-0.0156	-1.7222
2f ₄	—	—	—	0.2626
3f ₁	-0.7712	0.0445	0.4741	-4.8139
3f ₂	0.6153	0.8614	0.4805	9.4550
3f ₃	—	—	0.0902	-5.1928
3f ₄	—	—	—	0.9857
δ ₀₀₀	2.4968	0.9045	0.9072	3.4416
δ ₀₁₁	—	-1.0593	-1.0675	-1.0462
δ ₁₀₁	—	-0.2607	-0.2628	-0.2544
δ ₁₁₀	—	-1.0226	-1.0094	-0.9831

2,2), および3M(3,3,3,2) から3M(4,4,4,2) にすると, あまり明確ではないが適合度が改善された。さらに, 取り扱った水流量は渇水頻度をみるための月降水量であるため, 原点付近と確率変数の軸に沿う方向について推定分布とデータの適合度をみると, 3M(2,2,2,2) で実用的な精度で一致していた。

以上より, この例では, 3M(2,2,2,2) で一応満足される適合度が得られたと判断し, とがりの情報を持つ4次モーメントを導入すると, さらに適合度が改善される場合のあることがわかった。

ここで示した, 4.1,4.2 の2例のほかに, 3地点年最大日降水, 1地点年最大日流量とそれに対応した2地点の日降水についても, 3変数最大エントロピー分布を求めた。その結果, 上述の例と同様に3M(2,2,2,2) で一応満足される適合度が得られたと判断できた。

さらに次元を増やした4変数以上の取り扱い, 4重以上の積分を必要とするため, パラメタ同定のための計算時間はかなり大きくなる(たとえば, 東京大学大型計算機センターHITAC M-280Hで1時間以上)ことが予想されるが, 4重以上の数値積分も実行可能であり, 必要に応じて次元を増やしていくことができる。

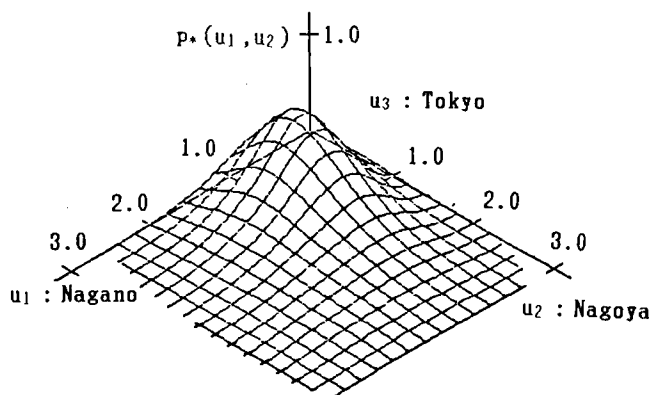
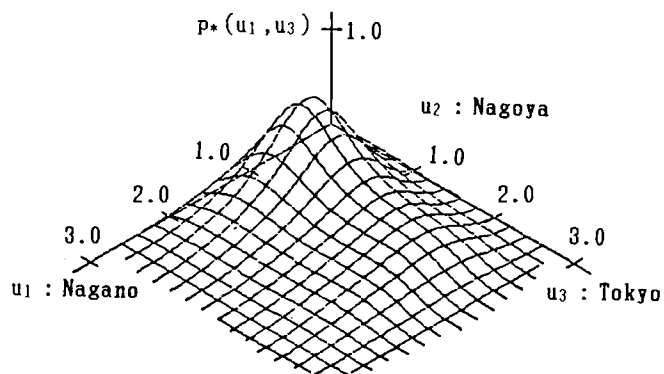
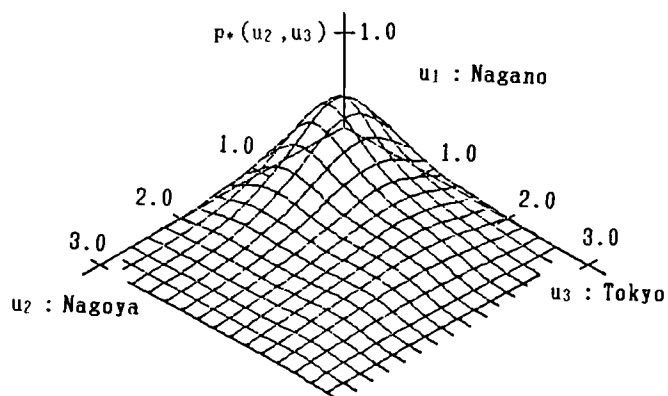


図4.5 3地点8月降水データより求めた $3M(4,4,4,2)$ の3次元立体図

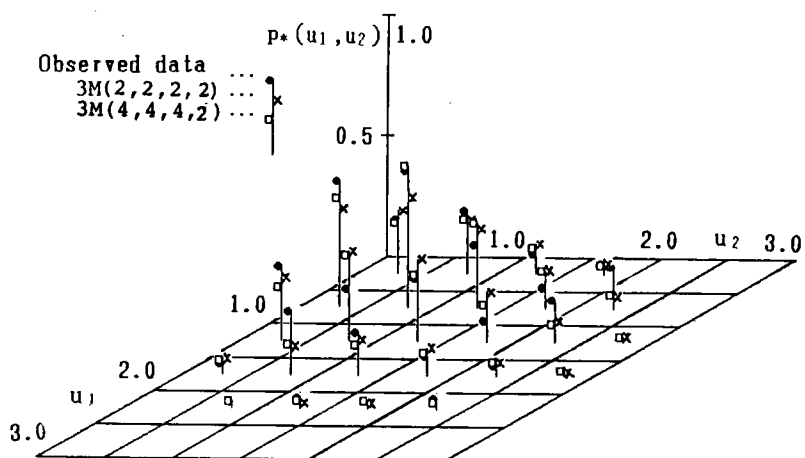
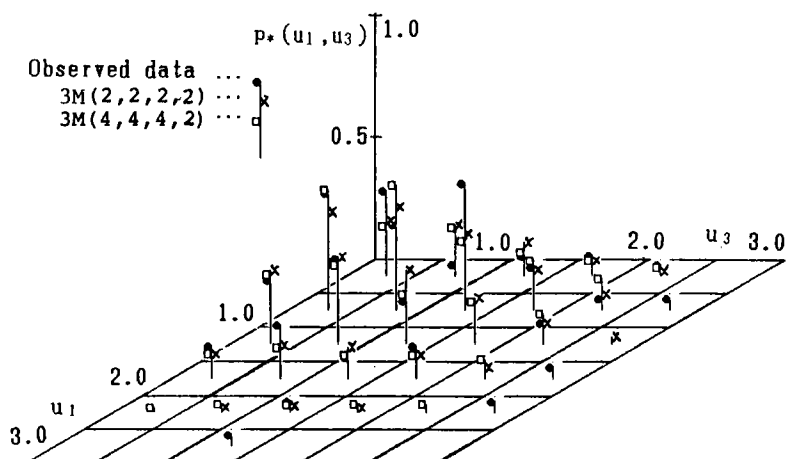
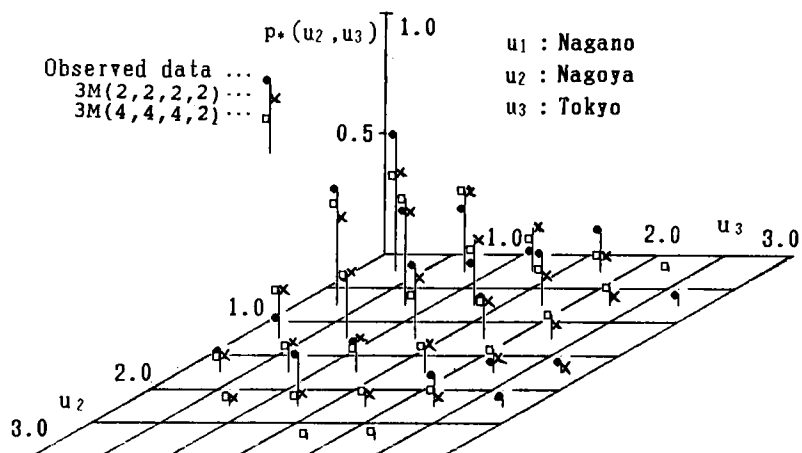


図4.6 $3M(2,2,2,2)$, $3M(4,4,4,2)$ およびデータのメッシュ確率

第5節 要 約

本章では、前章の2変数最大エントロピー分布を多変数の場合に拡張し、前章と同様な検討を行った。したがって、得られた成果については、2変数の場合から類推できることが多いが、以下にそれを要約する。

1) 2変数最大エントロピー分布を多変数に拡張して定式化し、そのパラメタ同定法を示した。また最大エントロピー分布から、既存の多変数分布の導出について触れておいた。

2) パラメタの値の組み合わせを変えた3つの3変数正規分布を母集団として、その真のモーメントを情報とした最大エントロピー分布で母集団推定を行った。その結果から、3変数正規分布のような母集団に対しては、相関係数に相当する2次の相乗モーメントを含む、すべての2次以下のモーメントを情報とすると、十分よい再現が得られることを示した。ただし、非対称形を示す指数分布以外の母集団を採用することができたなら、前章の結果から判断して、4次モーメントまで情報とすることが要求されよう。

3) 治水問題と渇水問題を念頭において、相関のある3地点の水文データに3変数最大エントロピー分布を適用した。この場合も、2)で述べたように2次モーメントまで情報とすると、全体的には一応満足される適合度が得られたが、やはりヒストグラムを精度よく記述するには4次モーメントまで情報とする必要があることを示した。

なお、ここでの計算は3変数の場合に留めたが、水工計画の策定において4個以上の水文量を同時に統計解析しなければならない場合が生じるであろう。このときは、本文中でも触れたように、当然4変数以上の最大エントロピー分布を用いることになるが、可能な変数の数の上限は、計算機が何重の数値積分まで対応できるかということと、パラメタ同定に要する実質的な計算時間の上限に依存してくることを付記しておきたい。

また、蛇足ではあるが、著者は、多変数分布が第1章で述べた現状にある以上、ここで定式化した分布は水文統計のみに留まらず、他の分野でも適用されんことを期待している。

参 考 文 献

- 1) 長沢敏夫：複合確率と河川工事計画における応用について、建設省直轄工事技術報告、第14回、pp.791~796、1960年。
- 2) 長尾正志：水文統計における多変数確率分布理論——二変数統計を中心として——、水工学に関する夏期研修会講義集(Aコース)、土木学会水理委員会、pp.A-4-1~A-4-19、1975年。

- 3) 石原安雄・長尾正志：計画高水流量の二変数確率論的研究，京都大学防災研究所年報，第6号，pp.181～188，1963年7月.
- 4) Krishnamoorthy, A. S. and Parthasarathy, M. : A Multivariate gamma-type distribution, The annals of mathematical statistics, Vol.22, pp.549 ~557, 1951.
- 5) 寒川典昭・荒木正夫・佐藤健次：多変数最大エントロピー分布とその基礎特性に関する研究，土木学会論文集，第375号/II-6，pp.89 ~98，1986年11月.
- 6) 例えば，笠原芳郎：情報理論と通信方式，共立出版，p.161，1965年10月.
- 7) 例えば，竹内 啓：数理統計学，東洋経済，p.47，1963年7月.
- 8) 例えば，前出7)，p.48.

第5章 条件付き最大エントロピー分布

第1節 概 説

本章では、前章で定式化した多変数最大エントロピー分布から、任意個数の変数を条件とし、残りの任意個数の変数を対象とした条件付き最大エントロピー分布を求め、実データへの適用をはかる。

一般に、水工計画を策定するとき、対象となる水文量を用いて、計画の重要度から設定されたリターン・ピリオドに対応する確率水文量が求められることは、前章までに述べてきたとおりであるが、統計解析を行う地点において、十分な量の水文データが整備されていない場合が多々ある。¹⁾ また、たとえ観測値が存在しても、その精度が著しく低く、統計解析に利用できないデータも見うけられる。¹⁾ このようなとき、近隣の観測所の精度の高い長期間の水文データから、不足する観測所の水文データを推定して、補充することが考えられるが、それを確率的に行うとき条件付き確率分布が必要となる。もちろん、水文統計において条件付き確率分布を要求される場面は上述の例のみでなく、長尾・角屋²⁾、端野ら^{3)、4)}の適用例にも見られるように多数存在している。

一方、理論面での条件付き確率分布としては、2変数正規分布⁵⁾、2変数 Gamma 分布（特別な場合として2変数指数分布を含む）^{6)、7)}、および Freund の2変数指数分布⁸⁾から導かれたもの以外は知られていないようである。したがって、適用では、当然のことながら、取り扱う水文量がこれらの分布に従うことが前提となるとともに、条件とする変数、および対象とする変数がともに1個の場合のみに限られることになる。

そこで、本章では、条件とする変数の数にも、対象とする変数の数にも、理論的には制限を加えない条件付き最大エントロピー分布を提案する。⁹⁾ 第2節では、この分布の定式化をはかるとともに、条件の数と制約方法を変化させながら、いくつかの導出した分布形を示す。第3節では、年降水量と年最大日降水量を用いて、この分布の実データへの適用を試みる。年降水量を使用した場合は、モーメントの取り方による分布形の変化と条件の取り方による分布形の変化を検討する。年最大日降水量を使用した場合は、対象とした地点の年最大日降水量を、それが条件とした地点との間で同時に発生することを仮定する方法と、模擬降水を利用する方法の両者から推定し、その結果を実測値と比較する。

なお、本章で述べる条件付き最大エントロピー分布は、前章までに記述した最大エントロピー分布の特徴を、特に小標本を取り扱う水文統計に応用するには有効な分布であり、密度関数

式が簡潔で利用しやすいという特徴を保持していることは言うまでもない。

第2節 確率密度関数の推定

2.1 定式化

ここでは、第4章2.1,2.3 の多変数最大エントロピー分布から条件つき最大エントロピー分布の定式化とモーメントを制約条件とした場合のパラメタ同定法について概説する。

一般に、確率変数 $x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n$ の同時確率密度関数を $p(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n)$ とするとき、 x_1, x_2, \dots, x_m を条件とした条件付き確率密度関数 $p_{x_1 x_2 \dots x_m}(x_{m+1}, x_{m+2}, \dots, x_n)$ は次式で与えられる。¹⁰⁾

$$\begin{aligned} & p_{x_1 x_2 \dots x_m}(x_{m+1}, x_{m+2}, \dots, x_n) \\ &= \frac{p(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n)}{\int \dots \int p(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n) dx_{m+1} dx_{m+2} \dots dx_n} \end{aligned} \quad (5.1)$$

したがって、式(4.5)を用いて、任意関数 $g_r(\cdot)$ を制約条件とした、 x_1, x_2, \dots, x_m の条件付き最大エントロピー分布は

$$\begin{aligned} & p_{x_1 x_2 \dots x_m}(x_{m+1}, x_{m+2}, \dots, x_n) \\ &= \frac{\exp \left\{ -1 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n) \right\}}{\int \dots \int \exp \left\{ -1 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n) \right\} dx_{m+1} dx_{m+2} \dots dx_n} \\ &= \frac{\exp \left\{ - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n) \right\}}{\int \dots \int \exp \left\{ - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_m, x_{m+1}, x_{m+2}, \dots, x_n) \right\} dx_{m+1} dx_{m+2} \dots dx_n} \end{aligned} \quad (5.2)$$

となる。

上式は一般的な多変数最大エントロピー分布であり，ここでも同様に種々の $g_r(\cdot)$ を与えることができる。

いま，統計モーメントで情報を与えることを考えると，制約条件は式 (4.14.1) ～ (4.14. n+1) となるため，このとき式 (5.2) は式 (4.15) を用いて次のように書き換えられる。

$$\begin{aligned}
 & P_{x_1 x_2 \cdots x_n} (x_{n+1}, x_{n+2}, \cdots, x_n) \\
 &= \exp \left(- \sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n} - \sum_{a_{n+1}=1}^{Na_{n+1}} \gamma_{a_{n+1}} x_{n+1}^{a_{n+1}} \right. \\
 &\quad \left. - \sum_{a_{n+2}=1}^{Na_{n+2}} \gamma_{a_{n+2}} x_{n+2}^{a_{n+2}} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n} - \sum_{b_n=0}^{Nb_n} \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \right) \\
 &\quad \left/ \int_0^\infty \cdots \int_0^\infty \int_0^\infty \exp \left(- \sum_{a_1=1}^{Na_1} \gamma_{a_1} x_1^{a_1} - \sum_{a_2=1}^{Na_2} \gamma_{a_2} x_2^{a_2} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n} \right. \right. \\
 &\quad \left. \left. - \sum_{a_{n+1}=1}^{Na_{n+1}} \gamma_{a_{n+1}} x_{n+1}^{a_{n+1}} - \sum_{a_{n+2}=1}^{Na_{n+2}} \gamma_{a_{n+2}} x_{n+2}^{a_{n+2}} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n} - \sum_{b_n=0}^{Nb_n} \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \right. \right. \\
 &\quad \left. \left. \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \right) dx_{n+1} dx_{n+2} \cdots dx_n \right. \\
 &= \exp \left(- \sum_{a_{n+1}=1}^{Na_{n+1}} \gamma_{a_{n+1}} x_{n+1}^{a_{n+1}} - \sum_{a_{n+2}=1}^{Na_{n+2}} \gamma_{a_{n+2}} x_{n+2}^{a_{n+2}} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n} - \sum_{b_n=0}^{Nb_n} \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \right. \\
 &\quad \left. \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \right) \\
 &\quad \left/ \int_0^\infty \cdots \int_0^\infty \int_0^\infty \exp \left(- \sum_{a_{n+1}=1}^{Na_{n+1}} \gamma_{a_{n+1}} x_{n+1}^{a_{n+1}} - \sum_{a_{n+2}=1}^{Na_{n+2}} \gamma_{a_{n+2}} x_{n+2}^{a_{n+2}} - \cdots - \sum_{a_n=1}^{Na_n} \gamma_{a_n} x_n^{a_n} \right. \right. \\
 &\quad \left. \left. - \sum_{b_n=0}^{Nb_n} \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \cdots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \right) dx_{n+1} dx_{n+2} \cdots dx_n \right. \quad (5.3)
 \end{aligned}$$

ここに， $\delta_{00 \cdots 0}$ の項は分子，分母間で消去されていると考える。その他のパラメタは第 4 章 2.3 の多変数最大エントロピー分布として同定された値を用いればよい。

次節では，式(5.3) で与えられる分布を ${}_n^2M(Na_1, Na_2, \dots, Na_n; Nb_1, Nb_2, \dots, Nb_n)$ と略記して使用する。

2.2 分布の導出

条件付き最大エントロピー分布も 1 変数，2 変数，多変数の場合と同じように形式的には無限個の分布形を記述することができるが，ここでは，原点まわりの統計モーメントを制約条件とした式(5.3) において，変数の数，条件の数，およびモーメントの次数を変えながら分母の積分を実行し，いくつかの導出した分布形を示す。

$$1) {}_1^2M(Na_1, 1; 0, 0)$$

$$p_{x_1}(x_2) = \frac{\exp(-2\gamma_1 x_2)}{\int_0^\infty \exp(-2\gamma_1 x_2) dx_2} = 2\gamma_1 \exp(-2\gamma_1 x_2) \quad (5.4)$$

ここに， $2\gamma_1 > 0$ とする。

$$2) {}_1^2M(Na_1, 1; 1, 1)$$

$$\begin{aligned} p_{x_1}(x_2) &= \frac{\exp(-2\gamma_1 x_2 - \delta_{11} x_1 x_2)}{\int_0^\infty \exp(-2\gamma_1 x_2 - \delta_{11} x_1 x_2) dx_2} \\ &= (2\gamma_1 + \delta_{11} x_1) \exp(-2\gamma_1 x_2 - \delta_{11} x_1 x_2) \end{aligned} \quad (5.5)$$

ここに， $x_1 \in [0, \infty)$ に対して $(2\gamma_1 + \delta_{11} x_1) > 0$ とする。

$$3) {}_1^2M(Na_1, 2; 1, 1)$$

$$\begin{aligned}
p_{x_1}(x_2) &= \frac{\exp(-2\gamma_1 x_2 - 2\gamma_2 x_2^2 - \delta_{11} x_1 x_2)}{\int_0^\infty \exp(-2\gamma_1 x_2 - 2\gamma_2 x_2^2 - \delta_{11} x_1 x_2) dx_2} \\
&= 2 \left(\frac{2\gamma_2}{\pi} \right)^{1/2} \frac{\exp\{-(2\gamma_1 + \delta_{11} x_1)^2 / (4 \cdot 2\gamma_2) - 2\gamma_1 x_2 - 2\gamma_2 x_2^2 - \delta_{11} x_1 x_2\}}{\operatorname{erfc}[(2\gamma_1 + \delta_{11} x_1) / \{2(2\gamma_2)^{1/2}\}]}
\end{aligned}
\tag{5.6}$$

ここに,

$$\operatorname{erfc}(z) = \frac{2}{\pi^{1/2}} \int_z^\infty \exp(-t^2) dt \tag{5.7}$$

であり, $2\gamma_2 > 0$ とする。

$$4) {}^2_1M(Na_1, 2; 2, 2)$$

$$\begin{aligned}
p_{x_1}(x_2) &= \frac{\exp(-2\gamma_1 x_2 - 2\gamma_2 x_2^2 - \delta_{11} x_1 x_2 - \delta_{12} x_1 x_2^2 - \delta_{21} x_1^2 x_2 - \delta_{22} x_1^2 x_2^2)}{\int_0^\infty \exp(-2\gamma_1 x_2 - 2\gamma_2 x_2^2 - \delta_{11} x_1 x_2 - \delta_{12} x_1 x_2^2 - \delta_{21} x_1^2 x_2 - \delta_{22} x_1^2 x_2^2) dx_2} \\
&= 2 \left(\frac{2\gamma_2 + \delta_{12} x_1 + \delta_{22} x_1^2}{\pi} \right)^{1/2} \cdot \exp\left\{-\frac{(2\gamma_1 + \delta_{11} x_1 + \delta_{21} x_1^2)^2}{4(2\gamma_2 + \delta_{12} x_1 + \delta_{22} x_1^2)} - 2\gamma_1 x_2 - 2\gamma_2 x_2^2 \right. \\
&\quad \left. - \delta_{11} x_1 x_2 - \delta_{12} x_1 x_2^2 - \delta_{21} x_1^2 x_2 - \delta_{22} x_1^2 x_2^2\right\} / \operatorname{erfc}\left\{\frac{2\gamma_1 + \delta_{11} x_1 + \delta_{12} x_1 + \delta_{22} x_1^2}{2(2\gamma_2 + \delta_{12} x_1 + \delta_{22} x_1^2)^{1/2}}\right\}
\end{aligned}
\tag{5.8}$$

ここに, $x_1 [0, \infty)$ に対して $(2\gamma_2 + \delta_{12} x_1 + \delta_{22} x_1^2) > 0$ とする。

$$5) {}^n_{n-1}M(Na_1, Na_2, \dots, Na_{n-1}, 2; \underset{\leftarrow n}{1}, 1, \dots, 1) \quad \rightarrow$$

$$\begin{aligned}
& p_{x_1 x_2 \dots x_{n-1}}(x_n) \\
&= \frac{\exp(-n\gamma_1 x_n - n\gamma_2 x_n^2 - \sum_{b_{n-1}=0}^1 \dots \sum_{b_2=0}^1 \sum_{b_1=0}^1 \delta_{b_1 b_2 \dots b_{n-1}} x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}} x_n)}{\int_0^\infty \exp(-n\gamma_1 x_n - n\gamma_2 x_n^2 - \sum_{b_{n-1}=0}^1 \dots \sum_{b_2=0}^1 \sum_{b_1=0}^1 \delta_{b_1 b_2 \dots b_{n-1}} x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}} x_n) dx_n} \\
&= 2 \left(\frac{n\gamma_2}{\pi} \right)^{1/2} \exp \left\{ - \left(n\gamma_1 + \sum_{b_{n-1}=0}^1 \dots \sum_{b_2=0}^1 \sum_{b_1=0}^1 \delta_{b_1 b_2 \dots b_{n-1}} x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}} \right)^2 \right. \\
&\quad \left. / (4 n\gamma_2) - n\gamma_1 x_n - n\gamma_2 x_n^2 - \sum_{b_{n-1}=0}^1 \dots \sum_{b_2=0}^1 \sum_{b_1=0}^1 \delta_{b_1 b_2 \dots b_{n-1}} x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}} x_n \right\} \\
&\quad / \operatorname{erfc} \left\{ \left(n\gamma_1 + \sum_{b_{n-1}=0}^1 \dots \sum_{b_2=0}^1 \sum_{b_1=0}^1 \delta_{b_1 b_2 \dots b_{n-1}} x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}} \right) / 2(n\gamma_2)^{1/2} \right\}
\end{aligned} \tag{5.9}$$

ここに, $n\gamma_2 > 0$ とする。

ここでは, 主要な5個の導出した分布形を示したが, この他にも式(5.3)の分母の積分可能性に依存して必要とされる分布形を導出することができる。

また, 第3章2.2で導出した式(3.16)の2変数正規分布から, 式(5.2)により次の条件付き正規分布が得られる。

$$p_{x_1}(x_2) = \frac{1}{(2\pi)^{1/2} \sigma_y (1-\rho^2)^{1/2}} \exp \left[-\frac{1}{2\sigma_y^2 (1-\rho^2)} \left\{ y - \mu_y - \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x) \right\}^2 \right] \tag{5.10}$$

第3節 実データへの適用

本節では, 水文データの推定, 発生のための基礎研究として, 実データから条件付き分布を求める。ただし, 相乗モーメントについては, 指定した最高次数までのすべての組み合わせを考慮するわけではないため, 2.1の略記方法を ${}_N M(Na_1, Na_2, \dots, Na_n; Nb)$ に変更する。ここに

N_{a1} は2.1 に同じであり、 N_b は採用した相乗モーメントの組み合わせの数である。

3.1 年降水量の場合

ここでは、図3.12に示した、長野、松本、上田、軽井沢の年降水量を用いた。年降水量は、原点付近にほとんど頻度がみられず、ヒストグラム全体は原点から遠く離れて存在することが多いため、パラメタ同定が困難な場合がある。したがって、本項での適用では、データから500mmを差し引いた値から基準化されたモーメントを計算してパラメタ同定を行い、図化するときにはもとの値にもどした。よって、確率変数の記述も基準化する前の形とした。また、計算にあたって、採用期間を次の3つに分割した。

- 1) Identification期間：条件とする地点のデータ、対象とする地点のデータが共にあり、分布の同定に利用する期間。
- 2) Checking期間：条件とする地点のデータ、対象とする地点のデータが共にあり、分布の同定に用いず、得られた結果の検討に利用する期間。
- 3) Estimation期間：条件とする地点のデータはあるが、対象とする地点のデータはなく、分布の推定に利用する期間。

3.1.1 モーメントの取り方による分布形の変化

表5.1 は採用した分布の記号の説明であり、表5.2 は条件とした地点、対象とした地点の組み合わせごとのデータ採用期間の表示である。また、表5.3 は得られた分布のパラメタの値である。いま、表5.3 のパラメタの値を用いて密度関数を描くと、モーメントの取り方により次のような分布形の変化を読み取ることができる。

- 1) 条件を松本 x_1 、対象を軽井沢 x_2 とした分布において、 x_2 に関するモーメントは2次まで用い、 x_1 に関するモーメントは、MK 1は0次（モーメントの採用なし）、MK 2は1次、MK 3は2次、MK 4は3次まで採用した。相乗モーメント $\int x_1^{b_1} x_2^{b_2} p(x_1, x_2) dx_1 dx_2$ を (b_1, b_2) と表現すれば、ここでは(0,0)、(1,1)の2通りを用いた。これらの分布形の1例を図5.1に示す。MK 1とMK 2は、分布の形状、期待値、標準偏差ともによく似ている。MK 2からMK 3にすると、分布形にシャープ度が加わり、標準偏差が小さくなるとともに、期待値も変動することがわかる。MK 3とMK 4は、MK 1とMK 2の場合と同様に、分布の形状のみならず、期待値、標準偏差も類似している。

ここで、MK 1～4は $\{N(N_{a1}, 2; 2)\}$ で表わされる分布である。したがって、理論式では条件 x_1 の N_{a1} 個のモーメントに対する N_{a1} 個のパラメタは、分子、分母間で消去されている

表5.1 年降水量から得られた最大エントロピー分布の記号の説明

($\prod M(Na_1, Na_2, \dots, Na_n; Nb)$ におけるNbは相乗モーメント

の数を表わす。他は2.1 に同じ。)

(a) $p_{x_1}(x_2)$

Symbol	Con- dition	Object	Distribution
MK 1	Matumoto (x_1)	Karuizawa (x_2)	$\frac{2}{1} M(0, 2; 2)$
MK 2	"	"	$\frac{2}{1} M(1, 2; 2)$
MK 3	"	"	$\frac{2}{1} M(2, 2; 2)$
MK 4	"	"	$\frac{2}{1} M(3, 2; 2)$
MK 5	"	"	$\frac{2}{1} M(2, 2; 5)$
NM 1	Nagano (x_1)	Matumoto (x_2)	$\frac{2}{1} M(2, 2; 2)$
NM 2	"	"	$\frac{2}{1} M(4, 4; 2)$

(b) $p_{x_1 x_2}(x_3)$

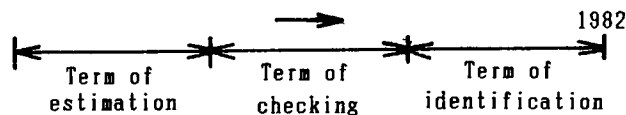
Symbol	Condition	Object	Distribution
NMK 1	Nagano, Matumoto (x_1) (x_2)	Karuizawa (x_3)	$\frac{3}{2} M(2, 2, 2; 2)$
NMK 2	"	"	$\frac{3}{2} M(2, 2, 2; 3)$
NUK 1	Nagano, Ueda (x_1) (x_2)	"	$\frac{3}{2} M(2, 2, 2; 2)$
NUK 2	"	"	$\frac{3}{2} M(2, 2, 2; 3)$
MUK 1	Matumoto, Ueda (x_1) (x_2)	"	$\frac{3}{2} M(2, 2, 2; 2)$
MUK 2	"	"	$\frac{3}{2} M(2, 2, 2; 3)$

(c) $p_{x_1 x_2 x_3}(x_4)$

Symbol	Condition	Object	Distribution
NMUK 1	Nagano, Matumoto, Ueda (x_1) (x_2) (x_3)	Karuizawa (x_4)	$\frac{4}{3} M(2, 2, 2, 2; 2)$

表5.2 年降水量データの採用期間

	Term of identification (Years)	Term of checking (Years)	Term of estimation (Years)
MK i ($i=1,2,\dots,5$)	40	17	28
NM i ($i=1,2$)	60	25	9
NMK i ($i=1,2$)	40	17	28
NUK i ($i=1,2$)	40	17	28
MUK i ($i=1,2$)	40	17	28
NMUK 1	40	17	18



が，パラメタ同定にはこのモーメントが影響を及ぼすため，条件付き分布の形状に上記のような相違が生じてきたものと判断される。

また，前章までの最大エントロピー分布では，偶数次数のモーメントを加えることにより，ヒストグラムに対する適合度の改善がみられたが，ここでも2次モーメントが加わると，分布形が大きく変化する様子がうかがわれる。

2) MK 5 は， x_1 ， x_2 のモーメントは2次まで用い，相乗モーメントを $(0,0)$ ， $(1,1)$ ， $(1,2)$ ， $(2,1)$ ， $(2,2)$ の5通りを採用した場合であり，その分布形の1例を図5.2 に示す。MK 5 は $\{N(2,2;5)\}$ であるから $\{N(2,2;2)\}$ のMK 3 と比較すると， x_1 の値が1000mm程度より小さいときMK 5 の方がシャープになり，1000mm程度以上になるとMK 3 の方がシャープ度を増す。図5.2 に示したように， x_1 の値が非常に小さいとき，MK 5 はかなりシャープな形状を示すことがわかる。

3) 条件を長野 x_1 ，対象を松本 x_2 とした分布において，NM 1 は x_1 ， x_2 に関するモーメントをそれぞれ2次まで用いた場合であり，NM 2 はそれぞれ4次まで用いた場合である。その分

表5.3 年降水量から得られた最大エントロピー分布のパラメタの値

(a) $p_{x_1}(x_2)$

	$2\gamma_1$	$2\gamma_2$	δ_{11}	δ_{12}	δ_{21}	δ_{22}
MK 1	-15.628	7.2780	0.94557			
MK 2	-13.070	6.8602	-0.64159			
MK 3	-7.4173	11.697	-15.968			
MK 4	-7.4153	11.695	-15.965			
MK 5	-29.850	24.760	24.837	-23.047	-17.072	9.4557

	$2\gamma_1$	$2\gamma_2$	$2\gamma_3$	$2\gamma_4$	δ_{11}
NM1	-4.0660	7.6915			-11.305
NM2	-1.2286	4.5097	1.1862	-0.054487	-11.381

(b) $p_{x_1x_2}(x_3)$

	$3\gamma_1$	$3\gamma_2$	δ_{011}	δ_{101}	δ_{111}
NMK 1	-14.103	9.0461			-3.8869
NMK 2	-3.1574	14.546	-17.014	-8.9181	
NUK 1	-16.082	9.8908			-3.5544
NUK 2	-8.9208	20.148	-22.456	-8.9176	
MUK 1	-16.908	10.534			-4.0041
MUK 2	-7.2122	23.341	-22.456	-17.013	

(c) $p_{x_1x_2x_3}(x_4)$

	$4\gamma_1$	$4\gamma_2$	δ_{1111}
NMUK 1	-14.449	7.2969	-0.72969

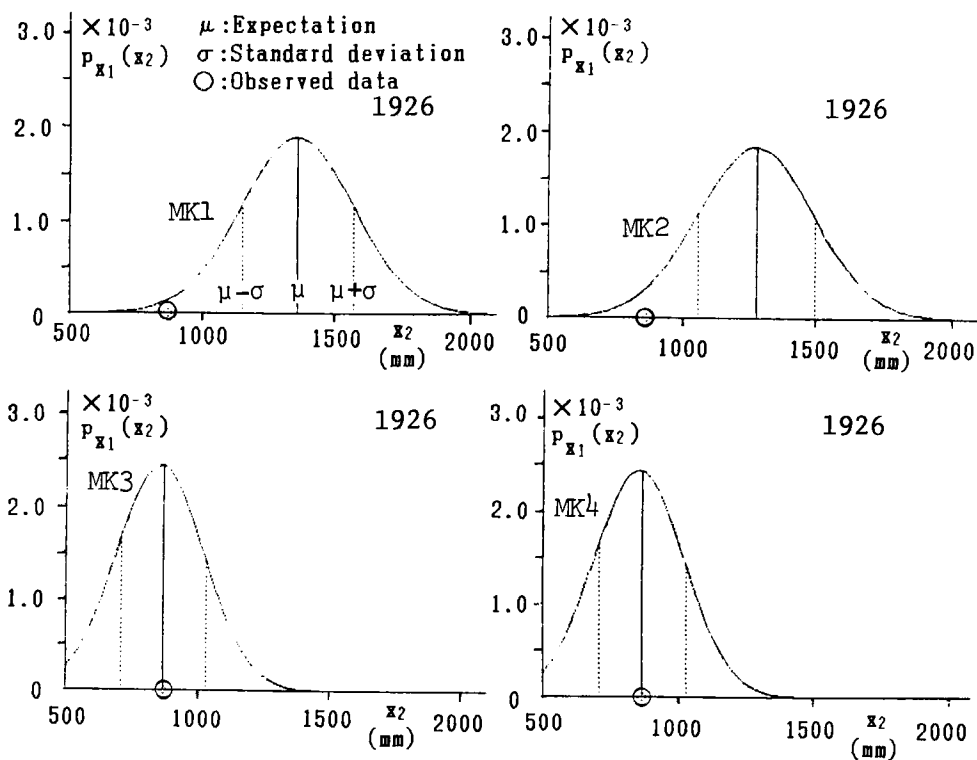


図5.1 モーメントの採用の仕方を変えた場合の分布形の変化
(x_1 :松本, x_2 :軽井沢, 相乗モーメント: (0,0), (1,1))

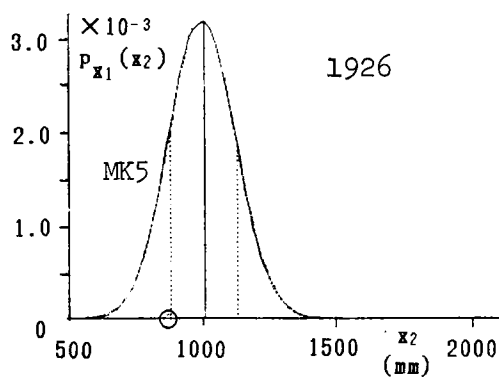


図5.2 モーメントの採用の仕方を変えた場合の分布形の変化
(x_1 :松本, x_2 :軽井沢, 相乗モーメント: (0,0), (1,1), (1,2), (2,1), (2,2))

布形の1例を図5.3に示す。ただし、相乗モーメントは(0,0)，(1,1)の2通り採用した。両者を比較すると、全体的によく似た形状を示しているが、 x_1 が1300mmを越えるところからNM2の分布の方がNM1よりシャープになっていることがわかる。

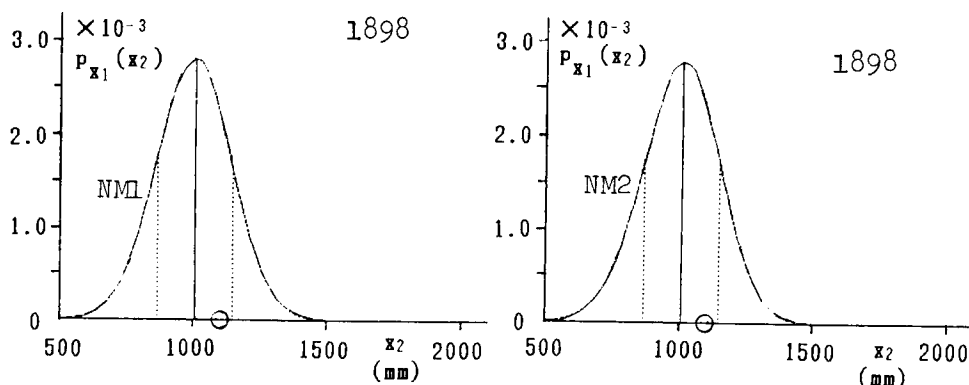


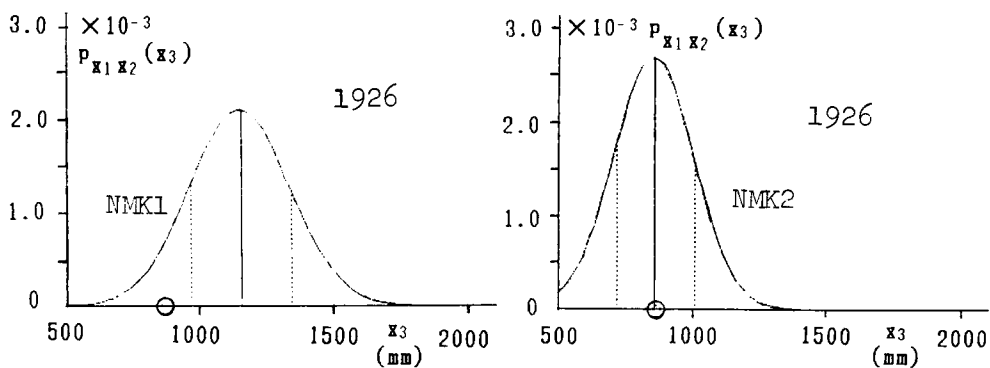
図5.3 モーメントの採用の仕方を変えた場合の分布形の変化
(x_1 :長野, x_2 :松本, 相乗モーメント:(0,0),(1,1))

4) 対象を軽井沢 x_3 とした分布において、NMK 1, 2は、条件を長野 x_1 , 松本 x_2 とした分布, NUK 1, 2は、条件を長野 x_1 , 上田 x_2 とした分布, MUK 1, 2は、条件を松本 x_1 , 上田 x_2 とした分布である。ここで、相乗モーメント $\int \int \int x_1^{b_1} x_2^{b_2} x_3^{b_3} p(x_1, x_2, x_3) dx_1 dx_2 dx_3$ を(b_1, b_2, b_3)と表現すると、NMK 1, NUK 1, およびMUK 1では(0,0,0), (1,1,1)の2通りを、NMK 2, NUK 2, およびMUK 2では(0,0,0), (1,0,1), (0,1,1)の3通りを採用した。これらの分布の1例を図5.4に示す。両者を比較すれば、どの場合も $\mathcal{M}(2,2,2;3)$ の方が $\mathcal{M}(2,2,2;2)$ よりはるかにシャープになっていることがわかる。

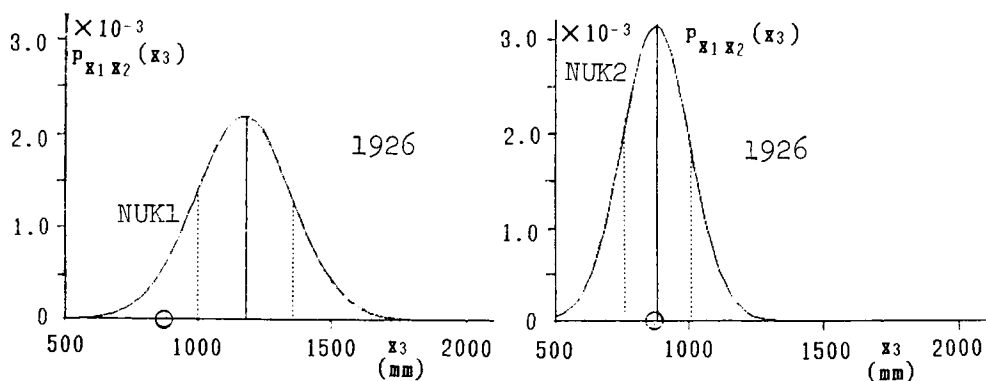
3.1.2 条件の取り方による分布形の変化

ここでは、条件を長野 x_1 , 松本 x_2 , 上田 x_3 , 対象を軽井沢 x_4 としたNMUK 1を追加した。これは、相乗モーメント $\int \int \int x_1^{b_1} x_2^{b_2} x_3^{b_3} x_4^{b_4} p(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4$ を(b_1, b_2, b_3, b_4)と表現すると、(0,0,0,0), (1,1,1,1)の2通りを採用した $\mathcal{M}(2,2,2,2;2)$ を表わしている。NMUK 1の分布の1例を図5.5に示す。なお、(0,0,0,0), (1,0,0,1), (0,1,0,1), (0,0,1,1)を用いた分布 $\mathcal{M}(2,2,2,2;4)$ は求まらなかった。

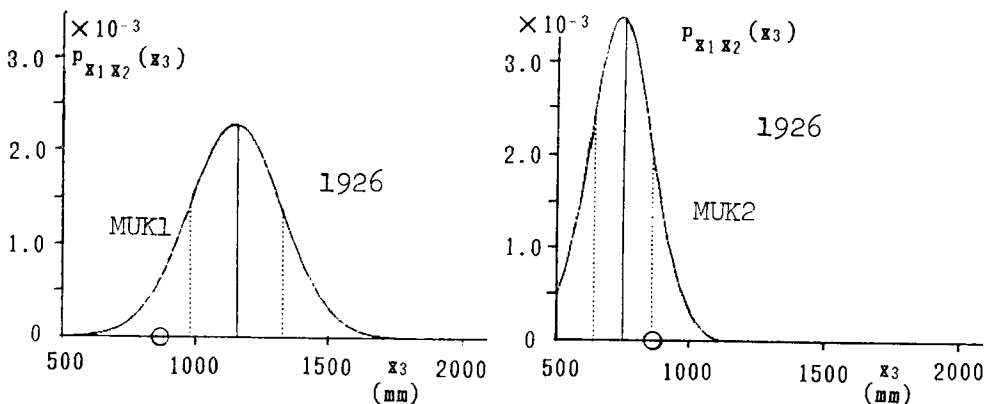
いま、条件数の変化による分布形の変動を調べるにあたり、対象を軽井沢、条件の1つを松本とした分布を用い、これらの分布を相乗モーメントの取り方により2つのグループに分けて考察する。



(a) x_1 :Nagano, x_2 :Matumoto, Product moments : NMK 1 =
(0,0,0) , (1,1,1) , NMK 2 = (0,0,0) , (1,0,1) , (0,1,1)



(b) x_1 :Nagano, x_2 :Ueda, Product moments : NUK 1 =
(0,0,0) , (1,1,1) , NUK 2 = (0,0,0) , (1,0,1) , (0,1,1)



(c) x_1 :Matumoto, x_2 :Ueda, Product moments : MUK 1 =
(0,0,0) , (1,1,1) , MUK 2 = (0,0,0) , (1,0,1) , (0,1,1)

図5.4 モーメントの採用の仕方を変えた場合の分布形の変化
(x_3 :軽井沢)

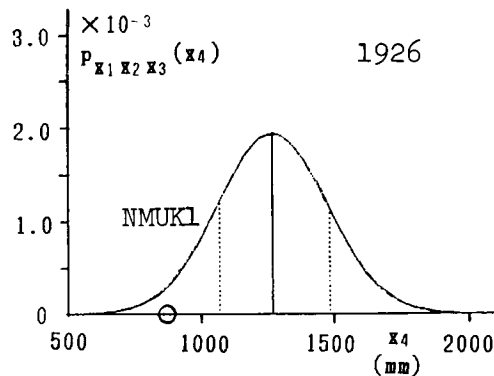


図5.5 条件の数を変化させた場合の分布形の変化 (x_1 :長野, x_2 :松本, x_3 :上田, x_4 :軽井沢, 相乗モーメント: $(0,0,0,0)$, $(1,1,1,1)$)

1) 第1グループ: MK 3, NMK 1, MUK 1, NMUK 1

このグループは, 相乗モーメントが, すべて0の組み合わせと, すべて1からなる分布である。この場合, 1条件のMK 3から, 2条件のNMK 1, MUK 1, 3条件のNMUK 1へと条件数が増加するに従い分布形状がフラットになる。これは, 相乗モーメントの取り方が適切でなかったため, 条件数の増加により分布形に平滑化作用が働いたためと推察される。

2) 第2グループ: MK 3, NMK 2, MUK 2

このグループは, 相乗モーメントがすべて0の組み合わせと, 対象とした軽井沢と条件のうちの1つが1になる組み合わせを用いた分布である。この場合, 1条件のMK 3から2条件のNMK 2, MUK 2へと条件数が増加すると, 分布形が明らかにシャープになる。これは, 変数間の相関に関する情報が, よりよく相乗モーメントに取り入れられたためと考えられる。また, MUK 2の方がNMK 2よりシャープになるのは, 軽井沢と上田の方が軽井沢と長野より相関が高いためと考えられる。

最後に, 参考までに, MK 3によって得られた期待値 μ と標準偏差 σ の値を, 軽井沢の実データとともに図5.6に示しておく。この図より, Identification期間からChecking期間に移っても, 推定精度はほとんど変化していないことがわかる。また, この傾向は他の分布からの推定値にも保たれている。

3.2 年最大日降水量の場合

本項では, 条件付き最大エントロピー分布を用いて, 長野のデータより松本の年最大日降水

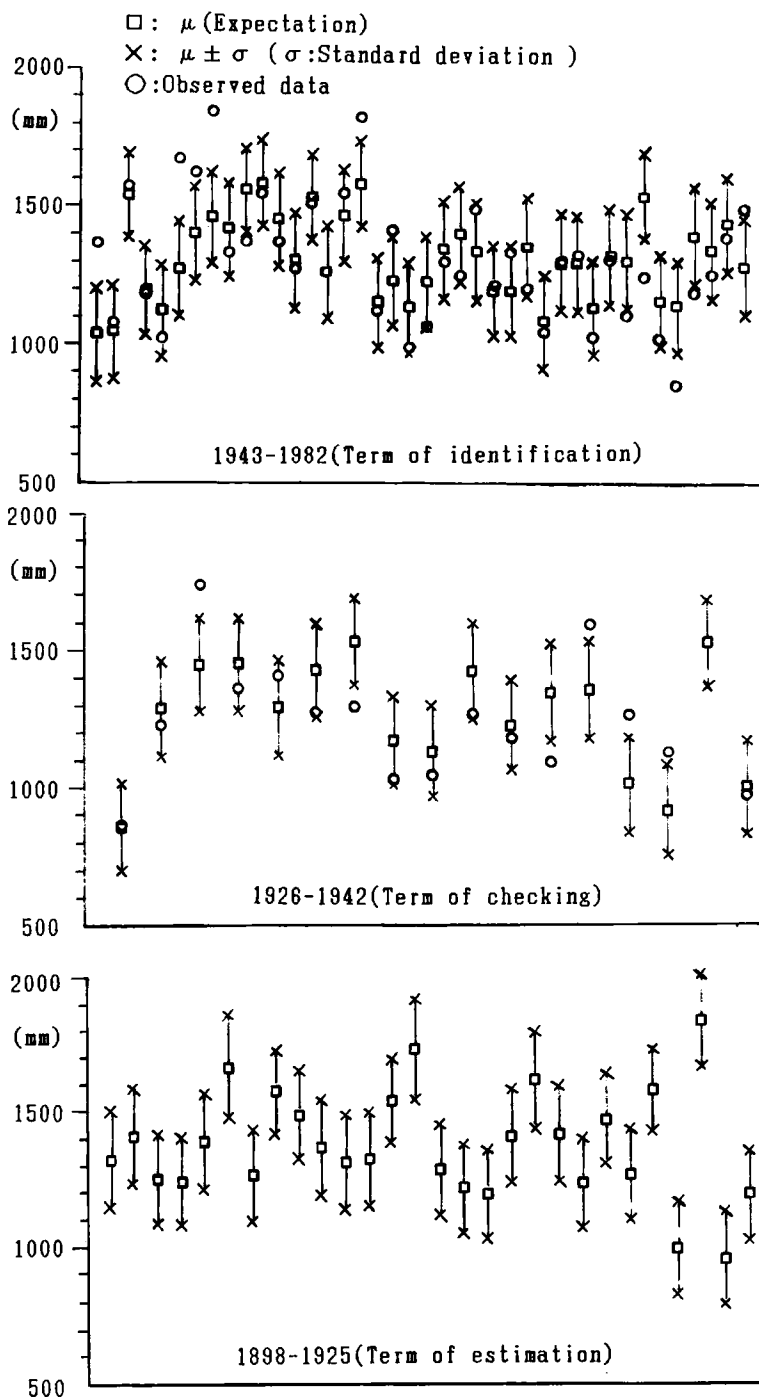


図5.6 MK 3 による年降水量の推定

量を推定する。なお、パラメタ同定には基準化したモーメントを用いたが図化するときにはもとの値にもどしたため、確率変数の記述は基準化する前の変数で行った。

3.2.1 同時発生を仮定した年最大日降水量の推定

ここでは、図3.12に示した長野と松本の年最大日降水量が同日に生起すると仮定して、長野年最大日降水量を条件 x_1 、同日の松本日降水量を対象 x_2 とした分布 $\{M(2,2;2)\}$ を求めた。ここで、相乗モーメントは $(0,0)$ 、 $(1,1)$ を採用し、Identification期間を1965年～1983年、Checking期間を1956年～1964年に設定している。

表5.4 は得られた分布のパラメタの値であり、図5.7 はChecking期間の4年分の分布形である。分布の形状はフラットな場合が多く、また、期待値と実測値の一致度はあまりよくないが、Identification期間からChecking期間になっても推定精度の低下はみられない。

図5.8 は、上述の $\{M(2,2;2)\}$ から得られた松本日降水量の期待値を松本年最大日降水量の推定値とみなして対数正規確率紙上に Weibull plot したものである。図5.9 は、図5.8 に対応した松本年最大日降水量の実測値の Weibull plot である。両者を比較すると、推定値の方が実測値より左にずれている。これは、長野と松本の年最大日降水が同じ日に生起するとした仮定の成立しない場合が多く、推定値が実測値より小さくなったためと考えられる。

3.2.2 模擬降水による年最大日降水量の推定

ここでは、長野における25mm以上の日降水量を条件 x_1 、同日の松本日降水量を対象 x_2 とした分布 $\{M(2,2;2)\}$ を求めた。ここで、相乗モーメントは $(0,0)$ 、 $(1,1)$ を採用している。また、Identification期間を1965年～1974年(147降水)、Checking期間を1956年～1964年(74降水)とした。

表5.5 は、得られた分布のパラメタの値であり、図5.10は、1959年における4降水の分布形である。

次に、 n 個のデータの最大値に関する Weibull plot による超過確率を $\alpha = 1/(n+1)$ とし、 $[\alpha, 1-\alpha]$ の上での一様乱数を使って、上述の分布から指定した x_1 に対応する x_2 を発生させた。²⁾ ただし、 α は1956年～1974年の221降水の計算から、0.005に設定している。

図5.11は、この模擬降水の年ごとの最大値を松本年最大日降水とみなして、対数正規確率紙上に Weibull plot したものである。図5.9 と図5.11を比較すると、後者の方がプロット点が左に広がっている。これは、長野の25mm以上の日降水と同日の松本の日降水との相関が低いため推定分布の形状がフラットになったこと、さらに対象とした長野の降水の年間発生個数が少なかったことに起因しているものと考えられる。

表5.4 年最大日降水量から得られた最大
エントロピー分布のパラメタの値

γ_1	γ_2	δ_{11}
3.0977	0.43992	-3.0490

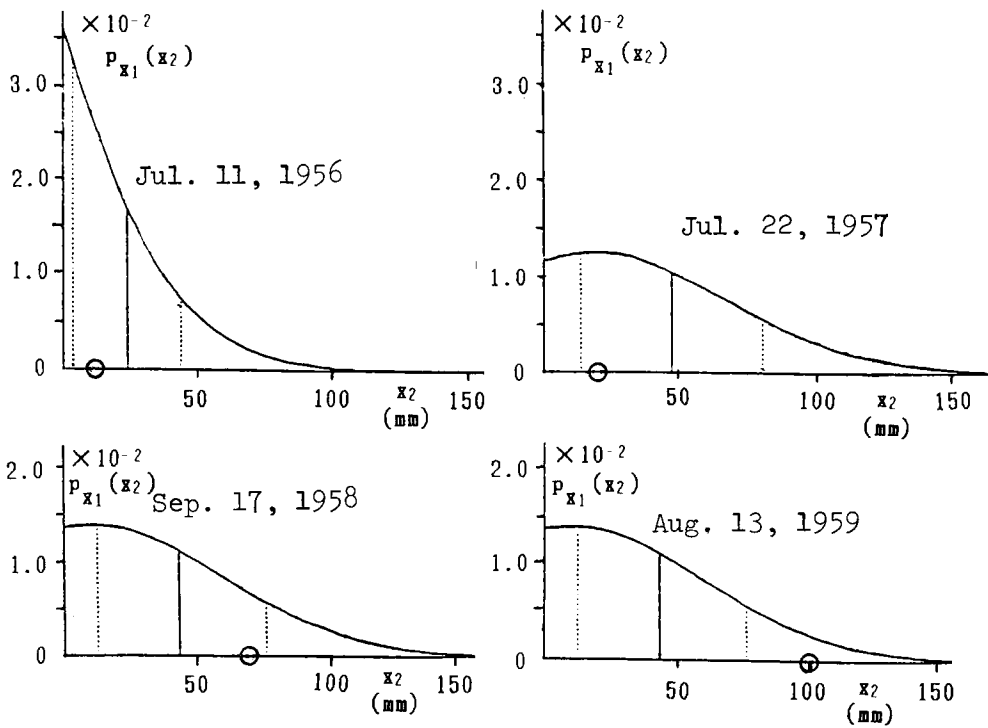


図5.7 年最大日降水量から得られた分布の形状 (x_1 :長野, x_2 :松本)

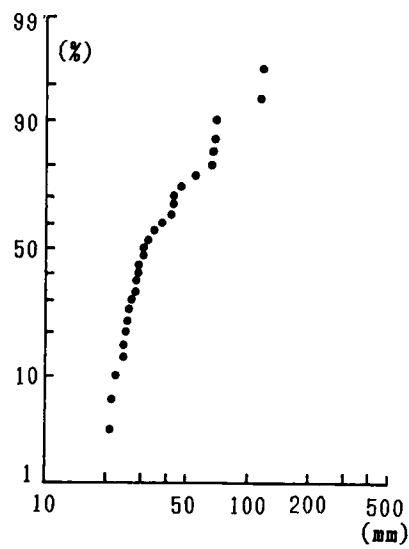


図5.8 年最大日降水量から得られた
推定値のWeibull plot (松本)

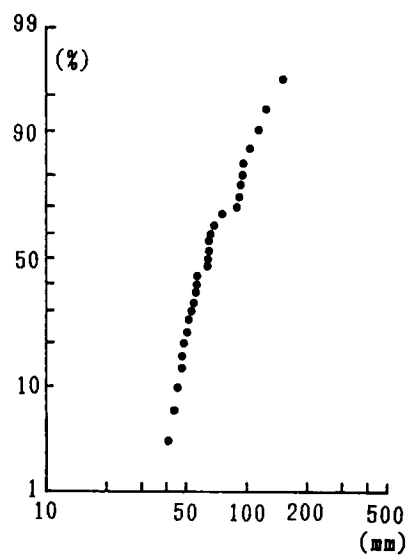


図5.9 年最大日降水量の実データ
のWeibull plot (松本)

表5.5 25mm以上の日降水量から得られた
最大エントロピー分布のパラメタの値

$2\gamma_1$	$2\gamma_2$	δ_{11}
-1.9152	0.35428	-1.8308

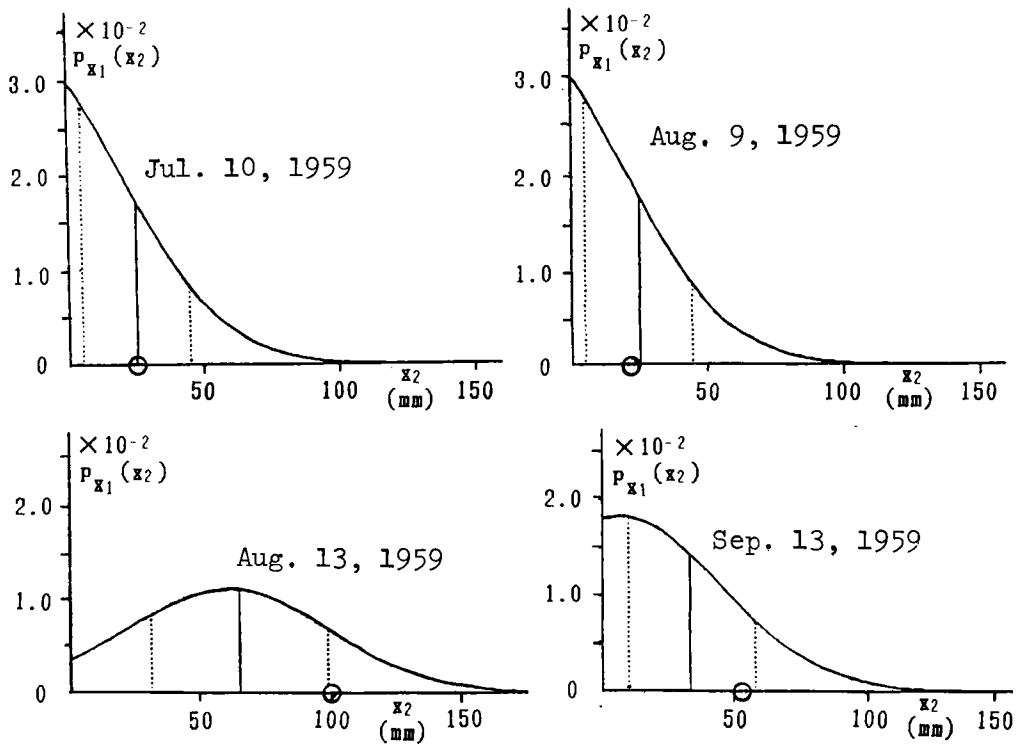


図5.10 25mm以上の日降水量から得られた分布の形状
(x_1 :長野, x_2 :松本)

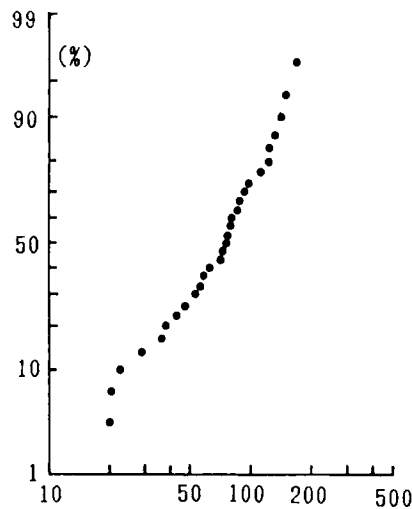


図5.11 シミュレーションデータから得られた
推定値のWeibull plot (松本)

第4節 要 約

本章では，前章の多変数最大エントロピー分布から，条件付き分布の定義に従って，条件付き最大エントロピー分布を導き，その特徴を検討した。以下に，得られた成果を要約する。

1) n 変数の最大エントロピー分布から m ($< n$) 個の変数を条件とし， $(n - m)$ 個の変数を対象とした条件付き最大エントロピー分布を定式化した。さらに，条件の数，および制約するモーメントの個数を指定して，5 個の条件付き最大エントロピー分布の理論式を導いた。また，この分布のパラメタ同定は，前章の方法に従えばよいことを述べておいた。

2) 条件付き最大エントロピー分布を年降水量に適用して，モーメントの取り方による分布形の変化と条件の取り方による分布形の変化を検討した。その結果から，特に相関構造の取り入れ方が条件付き分布の形状に大きく影響を及ぼすことを指摘した。

3) 次に，この分布を年最大日降水量に適用して，対象とした地点の降水量を，それが条件とした地点との間で同時に発生すると仮定する方法と，模擬降水を利用する方法の両者から推定した。このとき，推定値の Weibull plot を実測値のそれと比較すると，前者の方が小さな値が出る傾向にあることを示した。

なお，ここで，第2章～第4章で行ってきた仮定した母集団に対する適用性の検討を省略したのは，本章で用いた同定パラメタは，多変数最大エントロピー分布（2変数の場合を含む）として同定したものであるため，適合度も第3章，第4章の結果から類推できるからである。

参 考 文 献

- 1) 建設省北陸地方建設局千曲川工事々務所：千曲川・犀川河川要覧（雨量資料編），1976年 3 月。
- 2) 長尾正志・角屋 睦：二変数ガンマ分布とその適用に関する研究(4)，——短時間降雨の模擬発生——，京都大学防災研究所年報，第14号，B，pp.77 ～ 85，1971年 4月。
- 3) 端野道夫・杉 雄司：複数の2変数指数型分布の結合とその応用に関する研究，徳島大学工学部研究報告，第29号，pp.49 ～57，1984年。
- 4) 端野道夫：Marked point processに従う水文二変量の同時確率，第28回水理講演会論文集，pp.391～ 396，1984年 2月。
- 5) 例えば，長尾正志：水文統計解析，河川工学百年の歩みと淀川，p.88，1978年9 月。
- 6) 井沢竜夫：二変数の Γ -分布について（降水量の分布 第2報），気象と統計，第4 巻，第1号，pp.9～15，1953年。
- 7) 長尾正志：水文統計解析，河川工学百年の歩みと淀川，p.89，1978年 9月。
- 8) Freund, J. E. : A bivariate extension of the exponential distribution, American Statistical Association Journal, Vol.56, pp.971 ～977, Dec., 1961.
- 9) Sogawa, N., Araki, M. and Imai, T. : Studies on multivariate conditional maximum entropy distribution and its characteristics, Journal of Hydrosience and Hydraulic Engineering, Vol.4, No.1, pp.79 ～ 97, Apr., 1986.
- 10) 例えば，竹内 啓：数理統計学，東洋経済，p.44，1963年 7月。

第6章 情報の与え方と

最大エントロピー分布

第1節 概 説

前章までは、理論的には任意関数 $g_r(\cdot)$ で情報を与える形で記述したが、既知の母集団および実データへの適用にあたっては、その情報を分布の形状との関係が明らかにされている統計モーメントに限ってきた。本章では、できるだけ少ないパラメタから母集団を精度よく推定するために、パラメタ数を2個に固定して、適切な情報の与え方を検討する。^{1), 2)} ただし、議論を簡単にするため、本章では1変数の場合についてのみ考察する。

第2章 2.2では、いくつかの適した情報を与えることにより、1変数最大エントロピー分布から、統計学でよく知られている5個の1変数分布が導出されていることを紹介した。このことは、種々の情報の与え方を検討すれば、対象とする水文量ごとに適切な情報の与え方が存在することを示唆している。そこで、まず第2節では、任意関数の期待値を情報とした最大エントロピー分布のパラメタ同定法について記述する。第3節では、パラメタの組み合わせごとにGamma分布から6個の母集団を設定し、4つの代表的な関数の期待値を個々に情報とする場合について、シミュレーションデータを使って、母集団に対する最大エントロピー分布の適合度を検討する。さらに、この節では、前出の6個の母集団から代表的な2個の母集団を抽出し、上述の検討結果を踏まえて、べき関数と指数関数の期待値の組み合わせで情報を与えた最大エントロピー分布について、上述と同様な検討を行う。最後に、第4節では、前節で適合度のよかった後者の組み合わせた情報をもつ最大エントロピー分布を用い、5地点で得られた年最大日降水量のヒストグラムに対する適合性とリターン・ピリオドの安定性について考察する。

なお、本章では、最大エントロピー分布の“良さ”は、母集団の分布形、およびヒストグラムに対する全体的な形状の適合度を第1条件として、分布形の裾の部分（大きな値のリターン・ピリオド）の適合性、およびリターン・ピリオドの安定性を必要に応じ求め、判断している。

第2節 任意関数の期待値を情報とした分布のパラメタ同定法

第2章 2.1 は任意関数 $g_r(\cdot)$ の期待値で推定分布に情報を与えた一般的な場合の記述であ

るため、そこでの定式化はそのまま本節で採用することができるが、第2章 2.4 は $g_r(x)=x^r$ に限定しているので、ここでは、これを $g_r(\cdot)$ のままでパラメタ同定がなされ得るように理論式の拡張を測る。

いま、 $E[g_r(x)]$ を簡略して ν_r と書くと、式(2.3) は次のように書き換えられる。

$$\int g_r(x)p(x)dx = \nu_r, \quad r=1,2,\dots,N \quad (6.1)$$

そこで、式(2.6) を上式に代入し、式(2.2) を用いて整理すると、

$$\int g_r(x) \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x)\right\} dx = \nu_r \int \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x)\right\} dx, \quad r=1,2,\dots,N \quad (6.2)$$

なる非線形方程式が得られる。上式は、式(2.30)に対応した式であり、これから λ_r ($r=1, 2, \dots, N$) が決定される。 λ_0 は、決定された λ_r を用いて、式(2.6)を式(2.1)に代入して得られる次式から算定すればよい。

$$\lambda_0 = \ln \left[\int \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x)\right\} dx \right] \quad (6.3)$$

さて、ここでも非線形方程式(6.2) の未知数に関する線形化を第2章 2.4 に準じて行なうと、式(2.33) は

$$\sum_{j=1}^N (A_{i,j} - \nu_i A_{0,j}) \epsilon_j = A_{i,0} - \nu_i A_{0,0}, \quad i=1,2,\dots,N \quad (6.4)$$

に変更される。ここに、

$$A_{i,j} = \int g_i(x) g_j(x) \exp\left\{-\sum_{r=1}^N \alpha_r g_r(x)\right\} dx, \quad i=0,1,2,\dots, j=0,1,2,\dots \quad (6.5)$$

であり、 $g_0(x)=1$ とする。

式(6.4) は、式(2.33)の説明と同じ理由により解き得る。具体的には、第2章 2.4 で記述した1)～5)において、式(2.31)を式(6.3)に、式(2.33)を式(6.4)に読み換えればよ

い。

なお、本章では、情報の与え方と最大エントロピー分布の特性との関連を考察する基礎研究として1変数での検討を行なっているため、本節での記述もそれに準じたが、2変数以上への拡張も可能である。

第3節 既存分布に対する適用性

3.1 単一関数の期待値を情報とした場合

ここでは、 $g_r(\cdot)$ の具体的な関数形として、前章まで用いられてきた統計モーメント x^r を含み、 $1/x^r$ 、 $\exp(-rx/M)$ ($M: x$ の平均)、 $(\ln x)^r$ の代表的な4つの関数のそれぞれの期待値を個々に情報とする場合について検討を行った。なお、推定するパラメタの数と高次モーメントのもつ誤差をおさえるために、 $r=1, 2$ 、すなわち、各関数の2次までを情報とした。このとき、パラメタの数は、 λ_0 、 λ_1 、 λ_2 の3個であるが、 λ_0 は確率の総和が1であるという条件を満足するためのパラメタであり、 λ_1 、 λ_2 に依存して決定される。したがって、第2章2.1でも記述したように、ここで取り上げた分布は、2パラメタ分布と呼ぶべきであろう。

まず、母集団として多くの形状を表現できる分布が望ましいため、式(2.50)で表現されたGamma分布を用いた。今回は、指数分布 $G(0,1)$ 、および左にピークを持つ分布 $G(1,0.7)$ からしだいに右へピークを移動し、フラットになった $G(1,1)$ 、 $G(3,1)$ 、 $G(3,2)$ 、 $G(3,3)$ の6つを母集団とし、これより50個、100個、1000個の乱数を発生させ、そのシミュレーションデータを使って、式(6.1)から計算された情報に基づいて最大エントロピー分布を求めた。ここで、データ数50個、100個は現実に入手できる水文量の数の代表として取り上げており、データ数1000個はほぼ母集団に近い情報が得られた状態を想定している。したがって1000個の場合は既知の母集団に対する適合と同様に扱い得るものとみなして、以下の考察を行っている。図6.1、6.2の(a)は母集団分布とヒストグラムに対する推定分布の全体的な形状の適合度をみた1例であり、表6.1、6.2はそのときのパラメタの値である。

母集団 $G(0,1)$ は指数分布であり、単調減少している。この場合、部分的には $(\ln x)^r$ が良く適合している例があるが、全体的な形状を再現できるのは x^r の場合のみであった。特に乱数の数が1000個の場合、すなわち、ほぼ母集団に近い情報が得られた場合、 x^r で情報を与えた最大エントロピー分布は、すべての領域にわたってほとんど母集団に合致していた。また、データ数が1000個の場合、 $1/x^r$ で情報を与えた分布はパラメタの同定がなされなかった。

$G(0,1)$ 以外の他の母集団分布に対する適合では、どの数の乱数に対しても以下に述べた同様

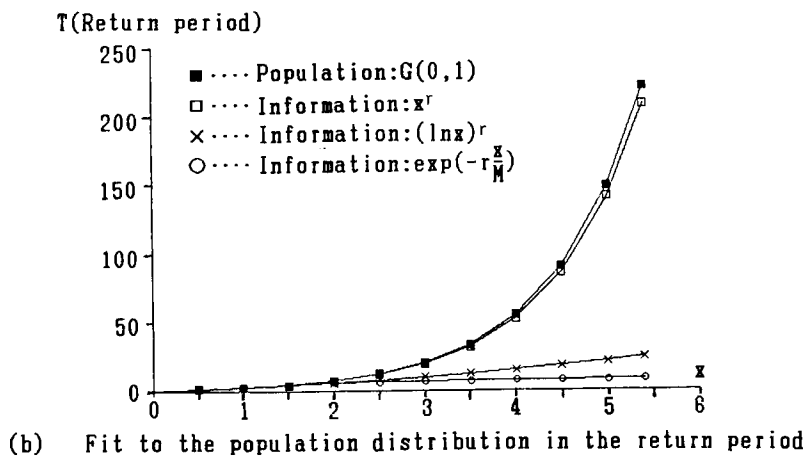
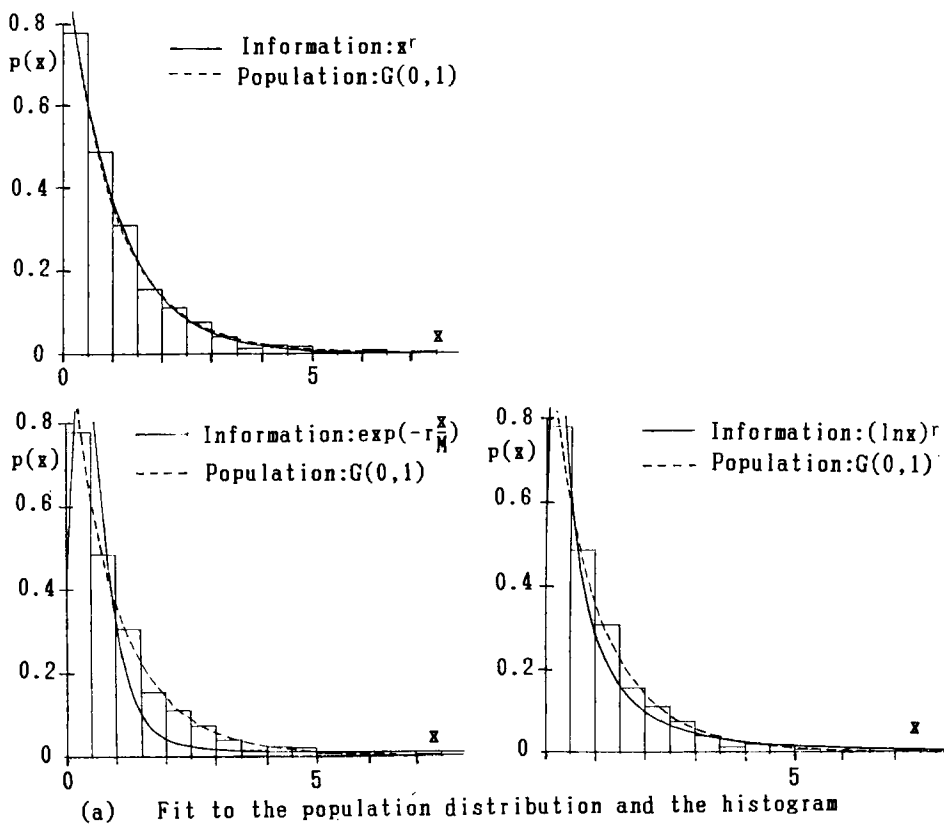


図8.1 $G(0,1)$ の最大エントロピー適合 (単一関数情報, データ数: 1000個)

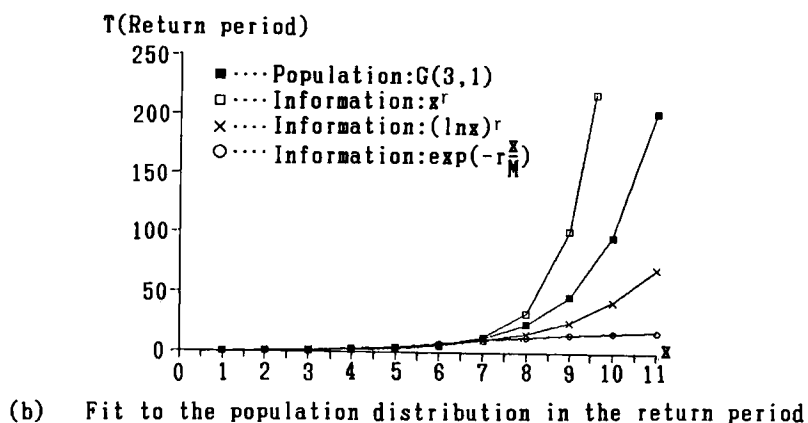
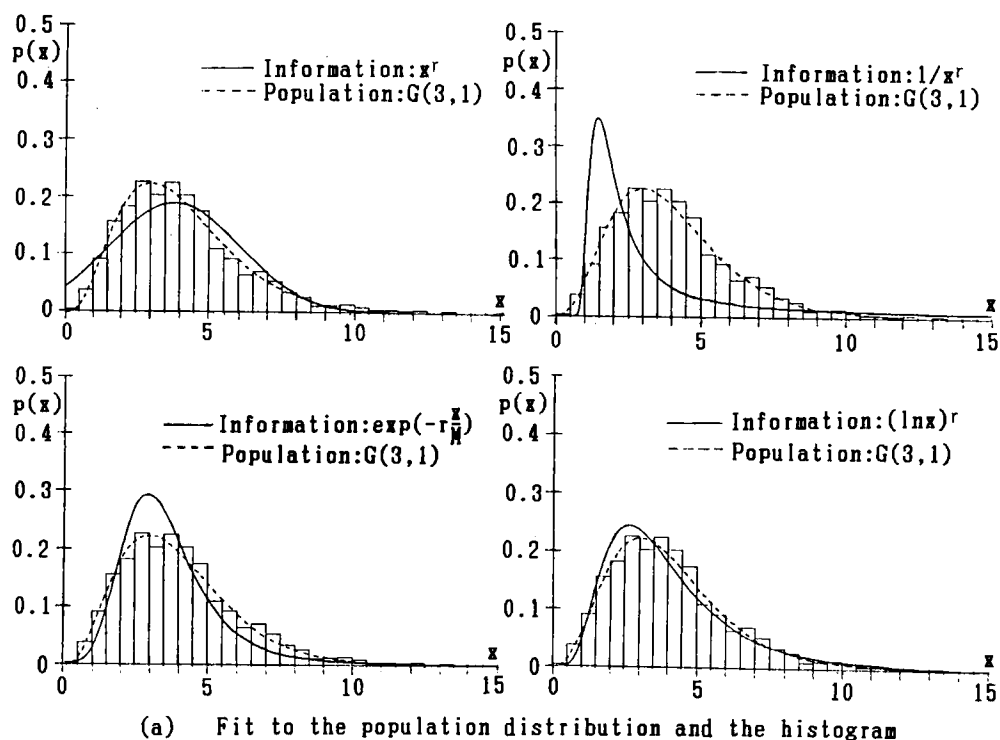


図6.2 $G(3,1)$ の最大エントロピー適合 (単一関数情報, データ数: 1000個)

リターン・ピリオドが、母集団よりかなり大きな値を示している。これは、100 個の場合のデータの分散が、50個および1000個の場合のデータの分散と比べて若干小さな値をとったことから、それに伴い推定分布も平均値付近に引かれ、右側にあまり尾を引かなかったためと推察される。また、他の母集団については、母集団の種類、乱数の数にかかわらず、 x の大きな値のところでのリターン・ピリオドは、 x^r の場合は母集団より大きく、他の2つの場合はかなり小さな値となっていることがわかる。

3.2 異種関数の期待値を情報とした場合

3.1 での結果から、 x^r と、 $(\ln x)^r$ か $\exp(-rx/M)$ のどちらか一方を組み合わせた分布は、単独の関数から得られた分布より全体的な適合度、及びリターン・ピリオドの点ですぐれていることが予想される。ところが、 $y=(\ln x)^r$ は r が奇数の時、負の領域を含み、 r が偶数の時、固定した y の値に対して2つの x の値が存在するため、他の関数と挙動を異にする。したがって、ここでは、 x^r と $\exp(-rx/M)$ の両方の期待値を情報として与える方法について検討する。

さて、この場合、最大エントロピー分布は、式(2.6) より次式のように表わされる。

$$p(x)=\exp \{-\lambda_0-\lambda_1 x^m-\lambda_2 \exp (-n x / M)\} \quad (6.6)$$

この分布の特性を知るために、上式中の (m, n) の値を変化させながら $G(1,1)$ と $G(3,1)$ を母集団とした場合について3.1 と同様な検討を行った。ただし、 $n=1$ とした場合は、パラメタ同定がなされなかったため、 m, n の値を、 $(m, n)=(1,2),(1,3),(1,4),(2,2),(2,3),(2,4),(3,2),(3,3),(3,4),(4,2),(4,3),(4,4)$ の12ケースとした。

図6.3(a)は、母集団分布とヒストグラムに対する推定分布の全体的な形状の適合度を、(b) はリターン・ピリオドの適合度をみた1例であり、表6.3 はそのときのパラメタの値である。

一般に、分布形は m の値が大きくなると、ピークの位置が少しずつ右下に下がり、フラットになる。また、あまり右側に尾を引かなくなるため、リターン・ピリオドが大きくなっていく。逆に、 n の値を大きくしていくと、ピークの位置が少しずつ左上に上がり、とがって全体的に左側に移動してくるが、右側に長く尾を引く分布形となるため、リターン・ピリオドが小さくなる。しかし、 m の値が3, 4となると、 n の値にあまり関係なく、母集団分布やヒストグラムに対する適合度が悪くなり、 x の大きな値に対するリターン・ピリオドも常に真値より大きくなる。従って、高次モーメントのもつ誤差を考慮に入れると、 m の値を2以下にするのが望ましい。

以上の考察より、式(6.6) の分布の特性が明らかになったが、母集団分布に対するリター

表6.1 $G(0,1)$ を母集団とした場合の最大エントロピー分布のパラメタの値
(単一関数情報, データ数: 1000個)

Information	λ_0	λ_1	λ_2
x^r	0.011001	0.989252	-0.000101530
$\exp(-r \frac{x}{M})$	4.83496	-13.5522	9.57854
$(\ln x)^r$	1.23908	1.32828	3.13964

表6.2 $G(3,1)$ を母集団とした場合の最大エントロピー分布のパラメタの値
(単一関数情報, データ数: 1000個)

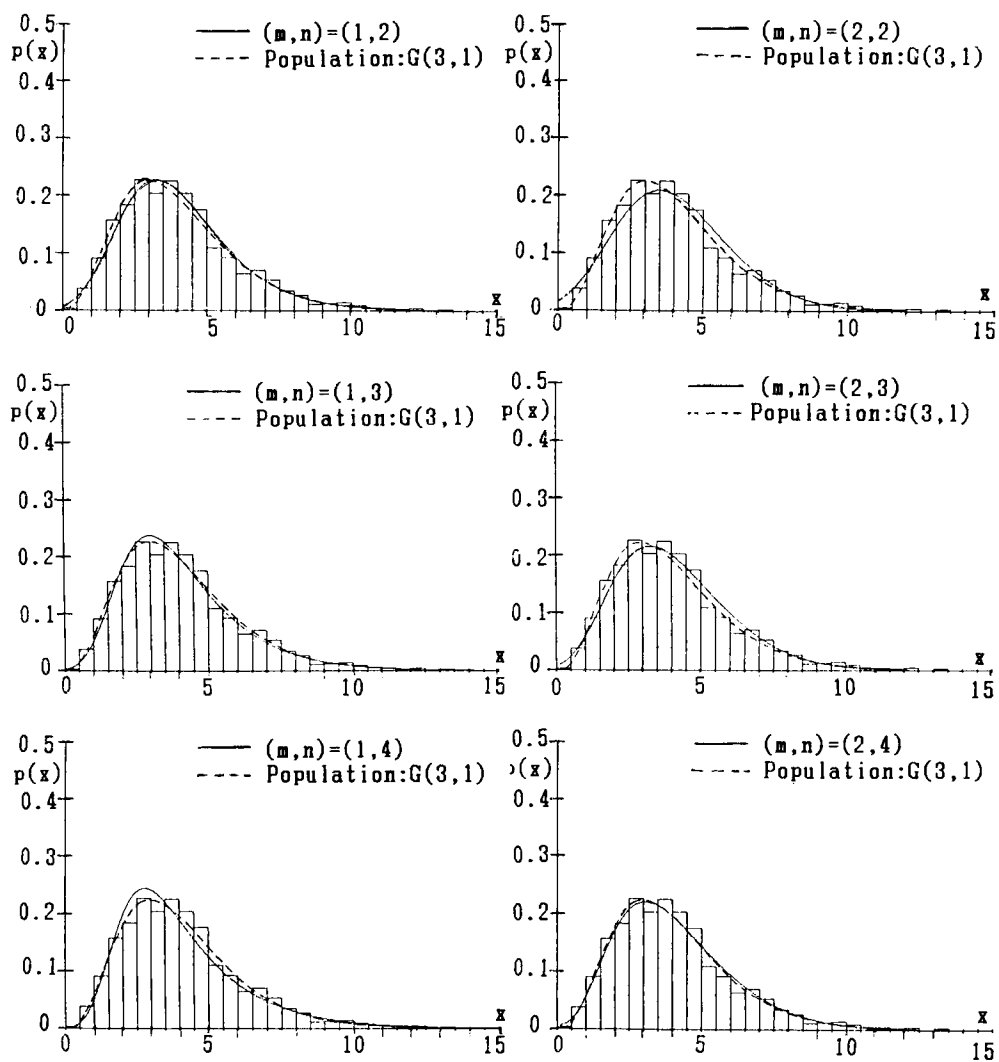
Information	λ_0	λ_1	λ_2
x^r	3.16682	-0.785051	0.102762
$\exp(-r \frac{x}{M})$	7.25759	-25.0539	26.0257
$(\ln x)^r$	3.05392	-3.38943	1.74394

な傾向を示した。

x^r の場合は, 他の関数で情報を与えた推定分布や母集団分布に比べてフラットな分布となり, 右側に長く尾を引かない。これは, r を2までとしたため, データから平均と分散の情報しか取り込まれなかったためと考えられる。逆に, $1/x^r$ の場合は, かなり全体的に左側に移動してとがった分布となり, 右側に長く尾を引いている。他の2つの場合は, 前の2つのほぼ中間的な分布となっており, $1/x^r$, $\exp(-rx/M)$, $(\ln x)^r$, x^r の順にピークが下がりフラットな形状を示すことがわかる。母集団に対する全体的な形状の適合度を比較すると, $(\ln x)^r$ のケースが一番すぐれており, その次に x^r と $\exp(-rx/M)$ が同程度で続き, $1/x^r$ が一番劣っている結果となっている。また, 分布形状から判断して, 母集団の左裾に対する適合では $\exp(-rx/M)$ と $(\ln x)^r$ のケースが, 右裾では x^r と $(\ln x)^r$ のケースがすぐれている。

そこで, 次に, x^r , $\exp(-rx/M)$, $(\ln x)^r$ の3つの場合について, 次節の極値水文量への適用において重要となる右裾の適合度をさらに詳細にみるために, 指定した x の値に対して求めたリターン・ピリオドを図6.1, 6.2の(b)に示した。

母集団を $G(0,1)$ とした場合, やはり x^r のケースの適合度が他に比べて良く, 上述の x^r のケースの説明と同様なことが言える。ただし, データ数が100個の場合のみ x の大きな値のところでの



(a) Fit to the population distribution and the histogram

図6.3 $G(3,1)$ の最大エントロピー適合 (異種関数情報, データ数: 1000個)

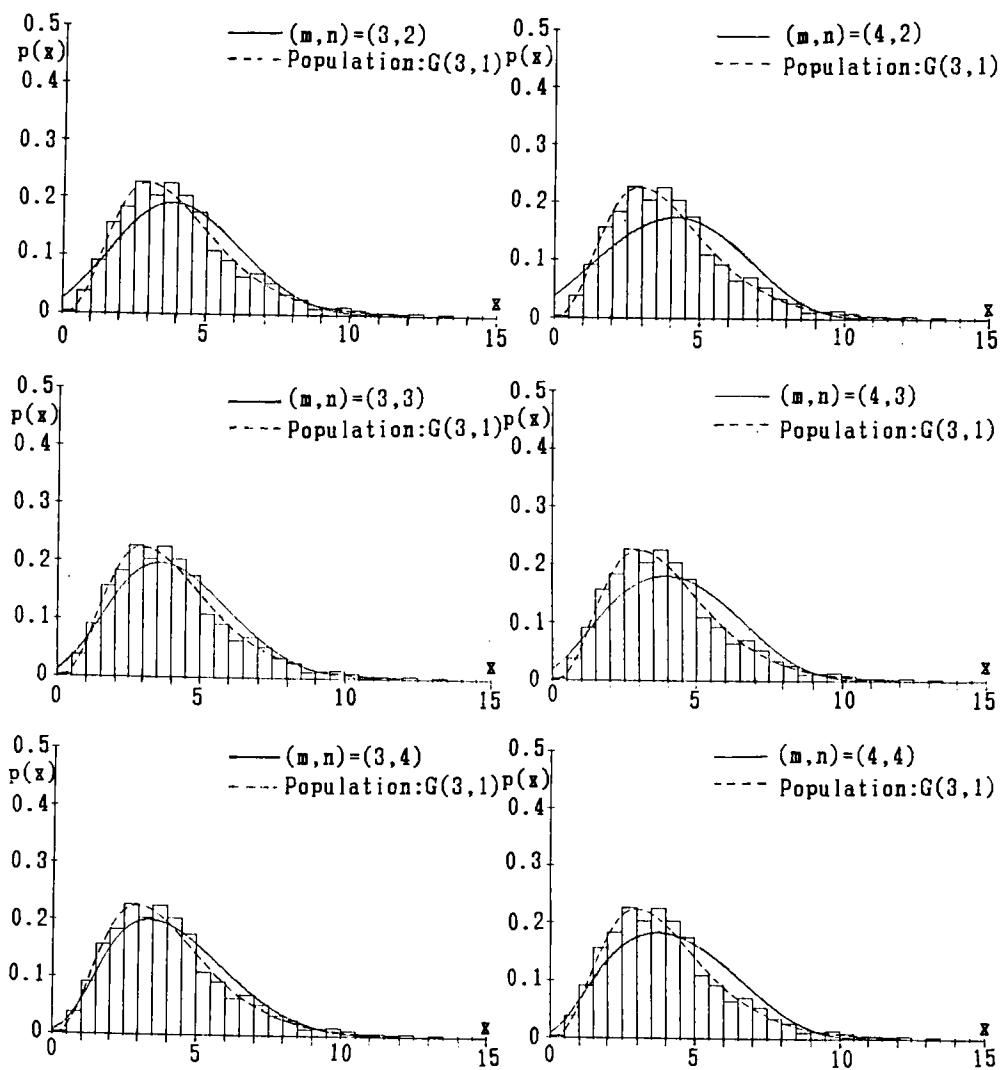
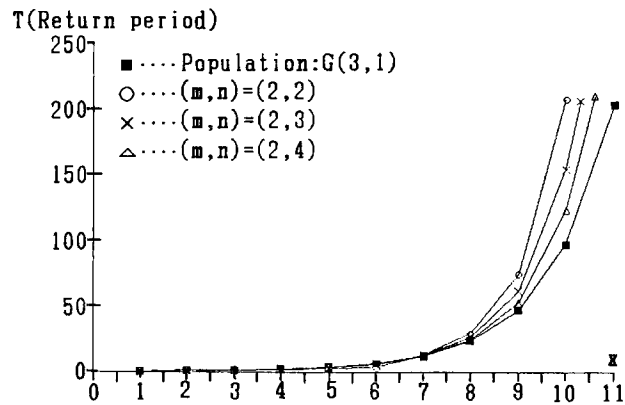
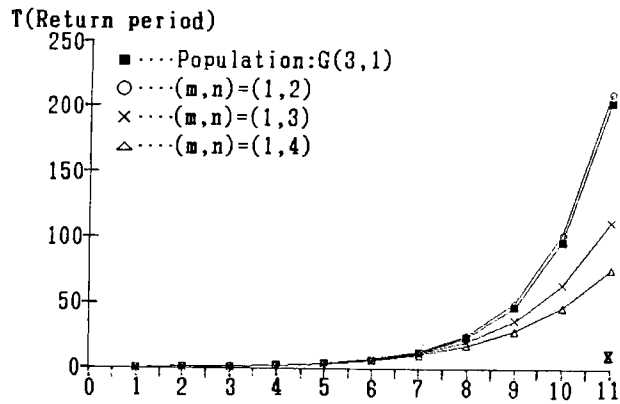


図6.3 (a) (続き)



(b) Fit to the population distribution in the return period

図6.3 $G(3,1)$ の最大エントロピー適合 (異種関数情報, データ数: 1000個)

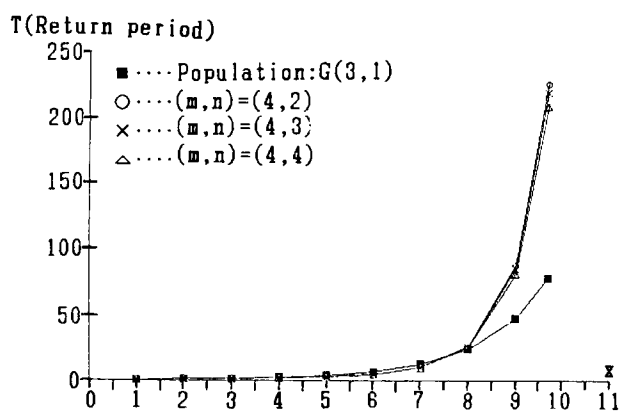
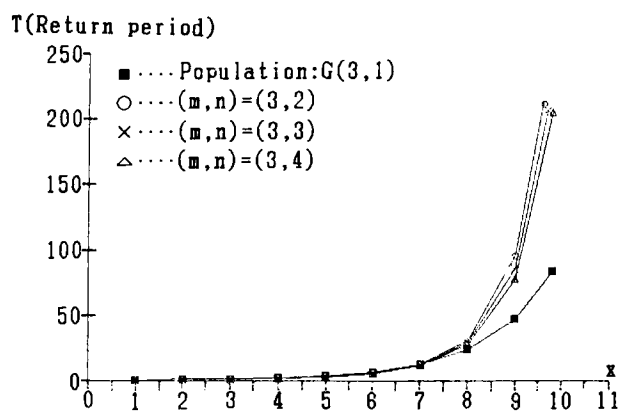


図6.3 (b) (続き)

表6.3 G(3,1)を母集団とした場合の最大エントロピー分布
のパラメタの値（異種関数情報，データ数：1000個）

(m,n)	λ_0	λ_1	λ_2
(1,2)	-2.32784	0.724318	7.33070
(1,3)	-1.09125	0.582312	7.32701
(1,4)	-0.51872	0.509197	8.14237
(2,2)	0.256135	0.0497977	3.97948
(2,3)	0.671831	0.0440867	4.44250
(2,4)	0.881155	0.0406055	5.26420
(3,2)	1.00786	0.00451533	2.64546
(3,3)	1.21056	0.00419823	3.19351
(3,4)	1.31780	0.00398059	3.96935
(4,2)	1.38474	0.000422198	1.89731
(4,3)	1.48766	0.000406042	2.48287
(4,4)	1.54746	0.000392253	3.22754

ン・ピリオドの適合性から判断して、G(1,1)のような左側にピークを持ち、とがった母集団に対しては、 $(m,n)=(1,2),(1,3),(1,4)$ を、またG(3,1)のようなフラットな母集団に対しては $(m,n)=(1,2),(1,3),(2,3),(2,4)$ を用いるとよい。

3.3 同数の異種データによる母集団推定

ここでは、母集団をG(1,1)とG(3,1)とし、これより50個と100個の乱数を10組ずつ発生させた。また、採用分布として、3.2で適合度の良かった、 $(m,n)=(1,2),(1,3),(1,4),(2,3),(2,4)$ を用い、これらの分布の、ヒストグラムと母集団分布に対する全体的な適合性、および次節の極値水文学に対する適用を念頭において、右裾の部分の挙動をみるために、母集団分布に対するリターン・ピリオドの適合性を検討した。

図6.4～6.7は、データ数が100個、母集団がG(3,1)の場合の、リターン・ピリオドの挙動をみた1例であり、表6.4は、G(3,1)および発生させたデータの平均・分散の値である。以下、表6.4の値に基づいて、図6.4～6.7を考察する。

図6.4は、平均、分散が共にほぼ母集団に等しい場合である。このとき、推定分布は、ヒストグラムに対してはもとより母集団分布に対しても非常によい適合度を示している。リターン・ピリオドについては、G(1,1)に対して、 $(m,n)=(1,3),(1,4)$ 、G(3,1)に対して、 $(m,n)=(1,2),(1,3),(2,3)$ の場合が、良く適合している。

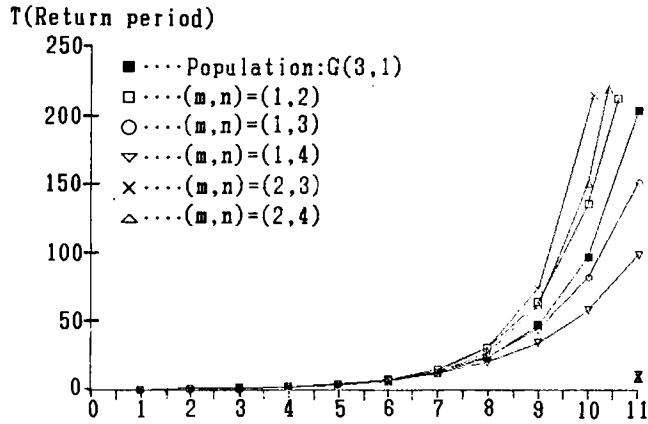


図6.4 $G(3,1)$ と平均，分散がともにほぼ等しい場合の
最大エントロピー適合（データ数：100 個）

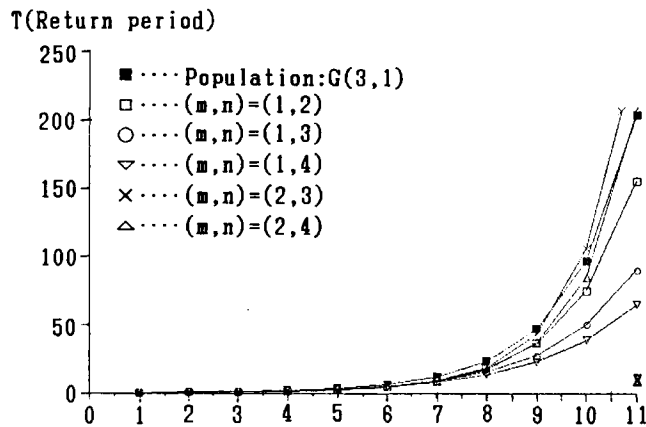


図6.5 $G(3,1)$ と分散がほぼ等しく，平均が大きい場合の
最大エントロピー適合（データ数：100 個）

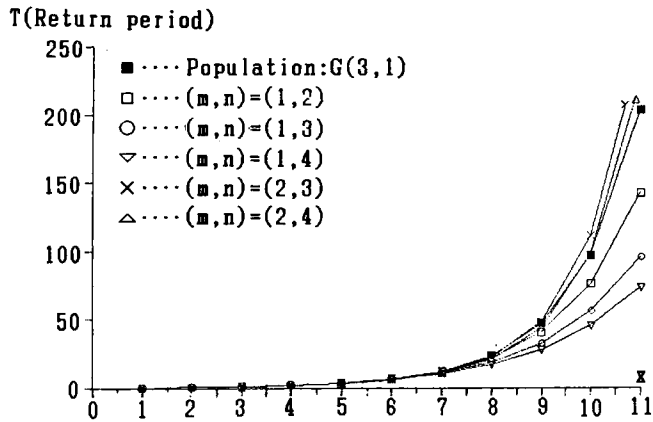


図6.6 $G(3,1)$ と平均がほぼ等しく，分散が大きい場合の
最大エントロピー適合（データ数：100 個）

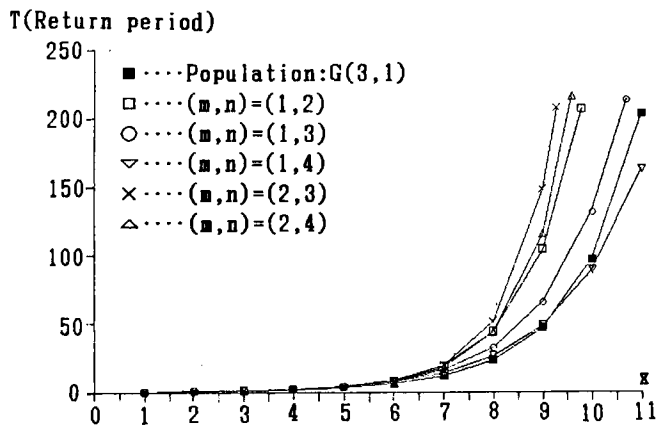


図6.7 $G(3,1)$ と平均がほぼ等しく，分散が小さい場合の
最大エントロピー適合（データ数：100 個）

表6.4 G(3,1)とシミュレーションデータの平均と分散

	Mean	Variance
G(3,1)	4.00	4.00
Fig.6.4	3.94	3.99
Fig.6.5	4.47	3.82
Fig.6.6	3.90	4.87
Fig.6.7	3.91	2.95

図6.5 は、分散が母集団とほぼ等しく、平均が母集団より大きい場合である。このとき分布形は全体的に母集団より右にずれている。このため、図6.4 に比べてリターン・ピリオドの値が全体的に小さくなり、G(1,1)に対しては $(m,n)=(1,2),(2,4)$ 、G(3,1) に対しては $(m,n)=(1,2),(2,3),(2,4)$ が良い適合度を示すようになっている。

図6.6 は、平均が母集団とほぼ等しく、分散が母集団より大きい場合である。このとき、 $(m,n)=(1,3),(1,4)$ は母集団よりとがり、 $(m,n)=(1,2),(2,3),(2,4)$ は、母集団よりもフラットになっている。リターン・ピリオドは、G(1,1)、G(3,1)のどちらの母集団に対しても、 $(m,n)=(1,2),(2,3),(2,4)$ が良く適合している。

図6.7 は、平均が母集団とほぼ等しく、分散が母集団より小さい場合である。この場合は、どの推定分布もとがった形状を示し、長く尾を引かなくなる。そのため、リターン・ピリオドの値が大きくなり、比較的長く尾を引く性質をもつ $(m,n)=(1,3),(1,4)$ が、G(1,1)、G(3,1)のどちらの母集団に対しても良く適合するようになってきている。

上述の考察は、乱数の個数が50個の場合についても同様であった。

以上より、最適な情報の与え方はある程度データに依存するが、採用する情報の候補を母集団ごとに少数個に絞ることができるとともに、ここで取り上げたすべての場合について、 $(m,n)=(1,3),(1,4)$ は比較的変動が小さく、小さな値のリターン・ピリオドをもつことがわかった。

第4節 実データへの適用

4.1 ヒストグラムに対する適用性

ここでは、3.2 でよい適合性を示した式(6.6) を、図3.12に掲載した、長野、上田、松本、

北牧，飯山の年最大日降水量に適用することを考える。なお， (m,n) の値は，前章の考察より， $(1,2), (1,3), (1,4), (2,3), (2,4)$ の5ケースを採用するとともに，第2章との比較のため， x^r ($r=1,2$) の期待値を情報とする方法での適用も行った。

図6.8 は，全データ数に対するヒストグラムと推定分布の1例であり，表6.5 はそのときのパラメタの値である。この図を見ると， x^r の場合の推定分布は，他の分布に比べてフラットで右側にピークを持ち，長く尾を引かない。また，他の分布の場合は，3.2 でも記述したように， m が大きくなるとピークが右下に下がってフラットになり， n が大きくなると全体的に左側に移動し，ピークが左上に上がってとがった分布形になっている。しかし，ここで用いたどの推定分布も，ヒストグラムに対してよい適合度を示していると言えよう。

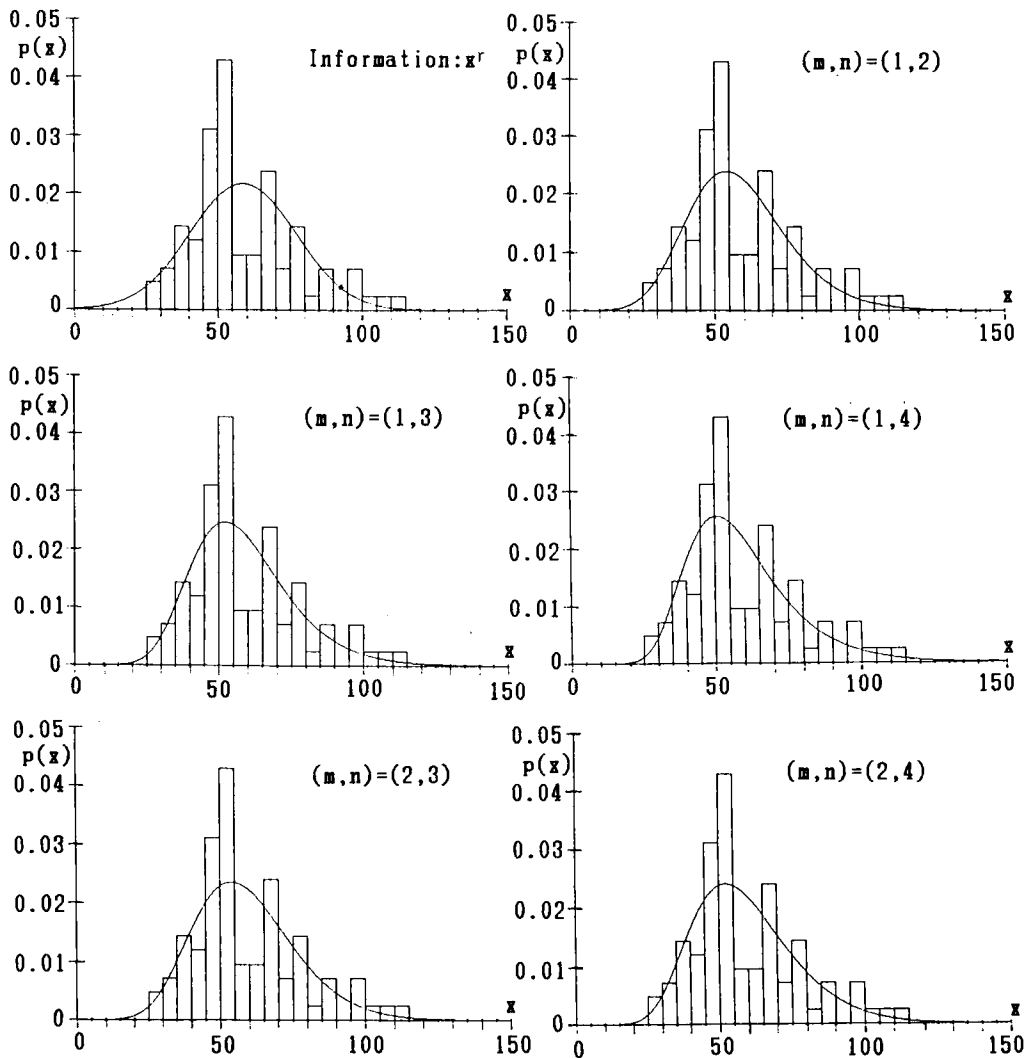


図6.8 長野年最大日降水量のヒストグラムへの適合性

表6.5 長野年最大日降水量から求めた最大エントロピー分布のパラメタの値

(m,n)	λ_0	λ_1	λ_2
x^r	8.91719	-1.73212	0.00147478
(1,2)	-5.44289	0.110106	20.3229
(1,3)	-2.27793	0.0832943	23.4049
(1,4)	-0.93203	0.0704374	32.4202
(2,3)	1.38295	0.000471869	15.6434
(2,4)	1.94040	0.000418127	22.6964

4.2 リターン・ピリオドの安定性

次に、上述のような極値水文量を取り扱う場合、分布形の右裾の方の特性を把握することが重要であるために、リターン・ピリオドの挙動について検討する。

図6.9 は、データを年代の古い順に並べ、前半分のデータの最大値のリターン・ピリオドが、前半分のデータに1つずつデータを追加することにより、どのような変動を示すかを図化した1例である。この図より、 x^r で情報を与えた場合は、どの地点においても、リターン・ピリオドの変動幅、値ともに大きいことがわかる。それに比較して他の場合はこれらの値が小さく、特に m が小さいほど、また n が大きいほど、その特徴が顕著にあらわれている。このことは、 x^r 以外の分布はかなり右側に長く尾を引き分布形状が比較的固定されているためと推察される。

一方、表6.6 は、全データ中の最大値のリターン・ピリオドを全データを用いて同定された最大エントロピー分布より求めたものである。 x^r の場合は、データ数、最大値、平均値から判断して、最大値のリターン・ピリオドとしては大きすぎる。他の場合は、上述したように、 m が小さく n が大きい程、リターン・ピリオドの値は小さくなっているが、飯山の場合を除いてはどの場合も実現可能性の高い値を示している。飯山では、全てのリターン・ピリオドが非常に大きな値をとっている。これは、データの平均値63.5mm/dayに対して、最大値156.5mm/dayが、きわだって大きいためである。しかし、 $(m,n)=(1,4)$ の場合は、639年となり、他に比べて小さい値に押えられていることがわかる。

以上より、式(6.6)の最大エントロピー分布は、年最大日降水量に対して、ヒストグラムとの適合度や、リターン・ピリオドの推定値および変動の考察から、かなり有効な分布であるこ

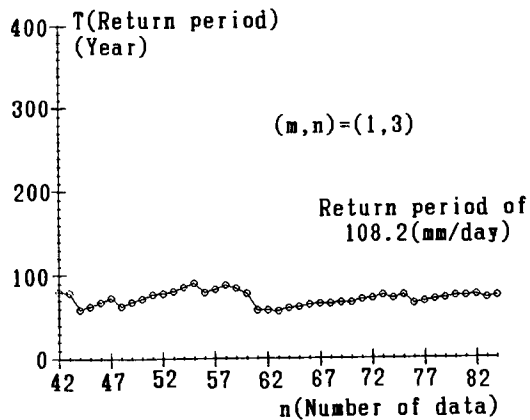
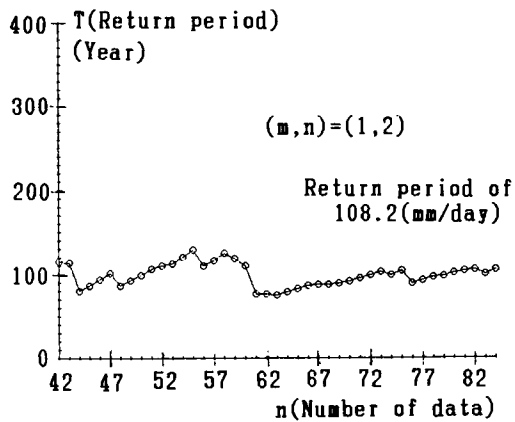
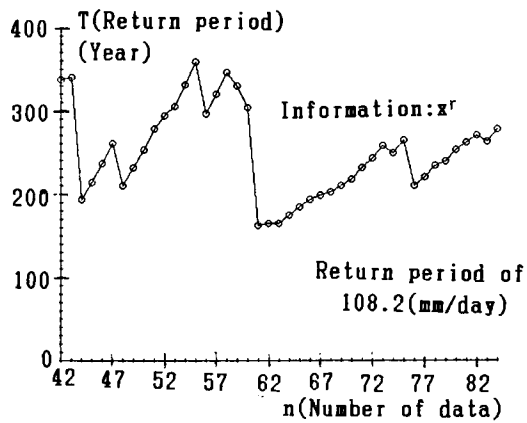


図6.9 リターン・ピリオドの安定性 (長野年最大日降水量)

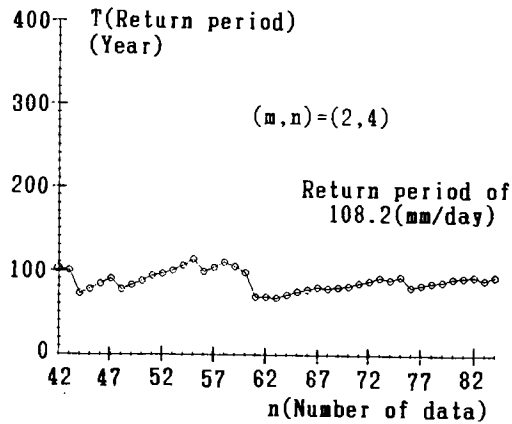
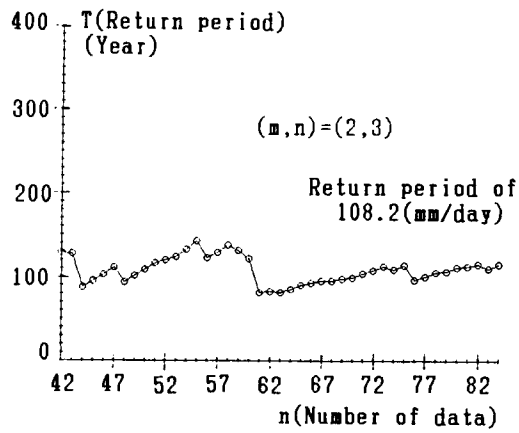
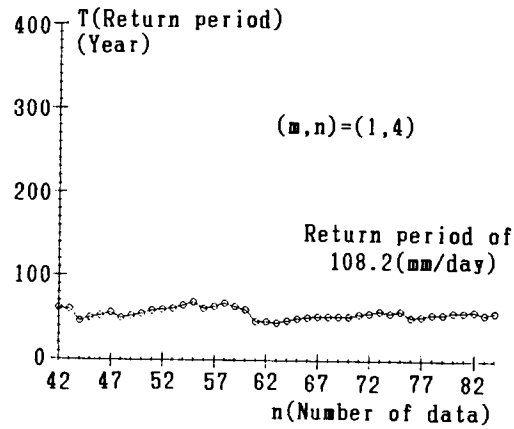


図16.9 (続き)

表6.6 最大値のリターン・ピリオド

Rain gauge station	Number of data	Maximum value (mm/day)	Mean value (mm/day)	x^r	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)
Nagano	84	113.9	58.8	730	183	118	85	216	163
Ueda	79	119.0	62.4	183	79	58	46	87	73
Matumoto	72	118.6	67.4	128	62	48	39	66	56
Kitamaki	54	134.0	75.2	86	39	30	25	43	36
Iiyama	77	156.5	63.5	100000	3430	1289	639	7181	4144

とが検証された。

ただし、 x^r で情報を与えた本節の結果は、さらに高次モーメントまで情報として利用することにより改善が期待されるが、ここでは他の分布とパラメタ数をそろえるため2次モーメントまでしか利用しなかったことを付記しておく。

第5節 要 約

本章では、パラメタの数を2個に固定して、母集団の分布形とヒストグラムに対する全体的な形状の適合度がよく、かつ分布形の右裾の部分の適合度にすぐれ、さらにリターン・ピリオドが安定した最大エントロピー分布を得るための情報の与え方を検討した。以下に、得られた成果を要約する。

1) 前章までは統計モーメントで情報を与えた場合のパラメタ同定法であったため、ここでは、それを任意関数の期待値を情報とした最大エントロピー分布のパラメタ同定法に拡張した。

2) まず、Gamma分布を母集団とし、情報を与える関数 $g_r(x)$ として x^r , $1/x^r$, $\exp(-rx/M)$ ($M=E[x]$), $(\ln x)^r$ を取り上げて、シミュレーションによる母集団推定を行った。この結果から、個々の関数を情報とする最大エントロピー分布では、母集団が指数分布のとき、 x^r の場合が、全体的な形状でも、リターン・ピリオドの点でもよい適合性を示すが、その他の母集団に対しては、特にリターン・ピリオドの点で、どの関数の場合も十分な精度の推定値を得ることができないことを述べた。

3) 次に、 x^n と $\exp(-nx/M)$ の2つの関数を組み合わせて情報を与えると、 $G(1,1)$ のような左側にピークを持ち、とがった母集団に対しては、 $(m,n)=(1,2),(1,3),(1,4)$ を、また $G(3,1)$ のようなフラットな母集団に対しては $(m,n)=(1,2),(1,3),(2,3),(2,4)$ を用いると母集団の分

布形に対する全体的な適合度のみならず、リターン・ピリオドの点でもよい適合度が得られることを示した。

4) さらに $G(1,1)$ と $G(3,1)$ の母集団から、50個、100個の乱数を10組ずつ発生させ、 $(m,n)=(1,2),(1,3),(1,4),(2,3),(2,4)$ で母集団推定を行い、ここで取り上げたすべてのデータの組について、 $(m,n)=(1,3),(1,4)$ は比較的変動が小さく、小さな値のリターン・ピリオドをもつことを示した。

5) 最後に、5地点で得られた年最大日降水量データを用いて、 $(m,n)=(1,2),(1,3),(1,4),(2,3),(2,4)$ で母集団推定を行って、どの分布もヒストグラムに対してよい適合度を示すとともに、リターン・ピリオドも比較的稳定して妥当な値を与えることを明らかにした。

なお、本章では、河川計画に最も重要な極値水文量に重点をおいて、適切な情報の与え方を考察したが、水文量が変われば5)とは異なる形の情報の方が良い場合が生じるであろう。例えば、対象とする水文量の母集団が指数分布に近い場合が、典型的な例である。しかし、最大エントロピー分布は、分布形を固定しないのがその特徴の1つであるから、その都度適切な情報の与え方を考えればよい。また、この分布は、分布形を固定しないかわり情報の与え方に主観が入り、それに分布形が依存して決定されるという点で、両者の間にトレードオフ的な意味合いが存在するが、この分布を用いると既存の分布よりすぐれた情報の与え方を探索することが可能であり、対象とする水文量（例えば年最大日降水量といった特定の時間単位の水文量）ごとに、情報の与え方を少数個に絞ることができるという点からも、最大エントロピー分布はかなり有効な分布であると言えよう。しかも、上述の利点を保持したまま容易に多変数に拡張され得るということは、従来の分布にない特徴である。ただ、我々はあくまでデータから得られる情報を通じて母集団を推定することが目的であるから、データにあまりに固執しすぎることは避けなければならない。

参 考 文 献

- 1) 荒木正夫・寒川典昭・大原基秀・森 茂：情報の与え方と最大エントロピー分布，土木学会中部支部研究発表会講演概要集，Ⅱ-4，pp.108～109，1987年3月。
- 2) 寒川典昭・荒木正夫・大原基秀：情報の与え方と最大エントロピー分布（その2），第42回土木学会年次学術講演会概要集，Ⅱ-3，pp.36～37，1987年9月。

第7章 最大エントロピー分布 の諸特性の検討

第1節 概 説

最大エントロピー分布は、既存の他の分布には見られない多くの特徴を有している。それらは、水文統計上重要なものもあれば、今のところ、それ程必要とされていない性質のものもある。しかし、現時点ではたとえ重要度が低くとも、この分布の多くの特徴を把握しておくことは、意義のあることと考えられる。

前章までに、必要に応じて最大エントロピー分布のいくつかの特徴を記述してきたが、本章では残された課題として、次の3点について検討する。¹⁾

- 1) 原点にパルス確率をもつ最大エントロピー分布とその適用性。
- 2) 高次モーメントまで制約条件とした最大エントロピー分布の適用性。
- 3) 複数の極値をもつヒストグラムへの最大エントロピー分布の対応。

また、最大エントロピー分布のパラメタ同定の基礎知識として必要であるため、Appendixとして、黒田・横田²⁾が原点のまわりの統計モーメントを制約条件とした離散的1変数最大エントロピー分布に対して行った、推定された確率密度関数がエントロピーを最大にすること、および解の一意性の証明が、多変数の任意関数を制約条件とした連続的多変数最大エントロピー分布についても同様に言えることを簡単に紹介する。¹⁾

なお、ここでの変数の数は、理論的には多変数で記述しているが、計算例では、視覚的に見やすくするために1変数の場合についてのみ取り上げた。

第2節 諸特性の検討

2.1 原点にパルスをもつ分布

日降水量のヒストグラムは原点に大きな頻度がみられる。ここでは、このような現象に対処するため、原点についてはパルスで確率を与え、その他の部分を最大エントロピー分布で適合させること考える。

2.1.1 定式化

原点にパルスを考えるため、 n 個の確率変数の存在範囲を $[0, \infty)$ とする。いま、 $p(0, 0, \dots, 0)$ を原点に与えたパルスの確率、 q をパルス以外の確率とすると、式 (4.2) は次の 2 式に書き換えられる。

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = 1 - p(0, 0, \dots, 0) \equiv q \quad (7.1)$$

$$\begin{aligned} \int_0^\infty \cdots \int_0^\infty \int_0^\infty g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \\ = E[g_r(x_1, x_2, \dots, x_n)], \quad r=1, 2, \dots, N \end{aligned} \quad (7.2)$$

ただし、式 (7.2) は $g_r(x_1, x_2, \dots, x_n) = 1$ を含まない。

このとき、最大エントロピー推定値は、式 (2.6) を参照すると

$$p(x_1, x_2, \dots, x_n) = \exp\left\{-\lambda_0 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \quad (7.3)$$

となることがわかる。

上式に含まれるパラメタは次のように推定される。

まず、式 (7.3) を式 (7.1) に代入して変形すると

$$\lambda_0 = \ln\left[\frac{1}{q} \int_0^\infty \cdots \int_0^\infty \int_0^\infty \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \cdots dx_n\right], \quad (7.4)$$

式 (7.3) を式 (7.2) に代入し、上式を用いて変形すると

$$\begin{aligned} \int_0^\infty \cdots \int_0^\infty \int_0^\infty g_r(x_1, x_2, \dots, x_n) \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \cdots dx_n \\ = \frac{E[g_r(x_1, x_2, \dots, x_n)]}{q} \int_0^\infty \cdots \int_0^\infty \int_0^\infty \exp\left\{-\sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \cdots dx_n \\ r=1, 2, \dots, N \end{aligned} \quad (7.5)$$

となる。したがって、 $\lambda_r (r=1, 2, \dots, N)$ は上式から決定され、 λ_0 は決定された λ_r を用いて、式 (7.4)から算定される。

式 (7.5)の非線形方程式は、1変数の場合、一般的には式 (7.5)の $E[g_r(x_1, x_2, \dots, x_n)]/q$ を式 (6.2)の ν_r に、統計モーメントで情報を与えた場合は式 (2.31) の μ_r に読み換えれば、そこで記述した方法により解かれ得る。統計モーメントで情報を与えた2変数の場合は、式 (3.28)～式 (3.30) の右辺を q で除せばよいことになる。多変数の場合は、式 (7.5)から離れて式 (7.1), (7.2)に式 (7.3)を代入した次の2式を考えた方が、第4章2.1 および2.3 の記述と整合性がとれて都合がよい。

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty \exp\{-\lambda_0 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\} dx_1 dx_2 \cdots dx_n = q \quad (7.6)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty g_r(x_1, x_2, \dots, x_n) \exp\{-\lambda_0 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\} dx_1 dx_2 \cdots dx_n$$

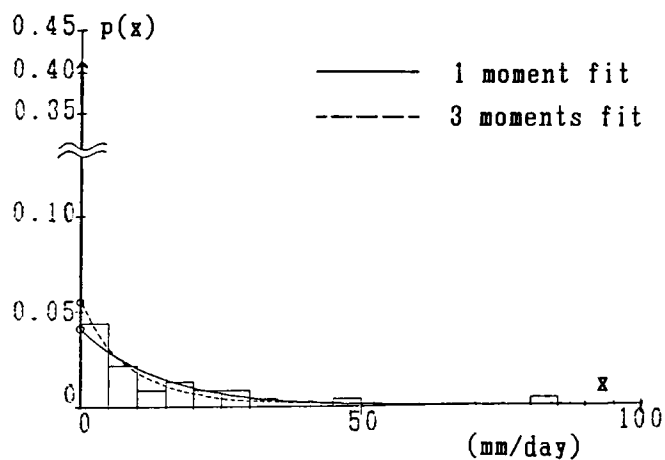
$$= E[g_r(x_1, x_2, \dots, x_n)], \quad r=1, 2, \dots, N \quad (7.7)$$

やはり統計モーメント制約を考えると、この場合も式 (4.14.n+1) の $\mu_{00 \dots 0}$ を q と置くことによって式 (4.14.1)～式 (4.14.n+1) の連立方程式を解く問題に帰すことができる。

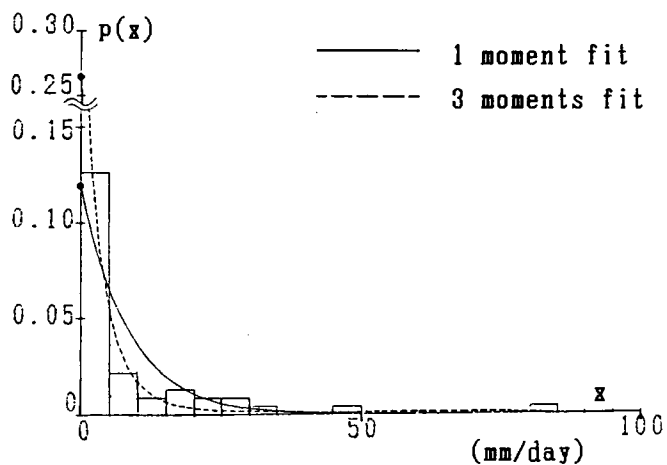
2.1.2 実データへの適用

解析データは、長野県の藪原(1938年～1983年、9月15日)と福島(1938年～1983年、6月全部)の日降水量とする。図7.1は藪原について、図7.2は福島について、モーメントを制約条件とした1変数最大エントロピー分布の適合度を、パルスを与えた場合と与えない場合で比較したものである。両図より、パルスを与えた場合原点付近の適合度の改善がみられ、その有効性が読み取れる。また、上述の期間での、藪原(6月全部)、福島(9月15日)についても同様な結果を得ている。

ただし、ここでの適用では、パルス確率の値は、実測値から計算された零となる確率を用いているため、パルスを与えた推定分布は、パルスを与えない推定分布より、データから得られた情報を1つ多く利用していることになる。したがって、情報の増加に伴い適合度の改善がなされたとみることもできるが、この場合、パルスという情報の与え方は、他の情報の与え方よりも効果的であると判断できる。

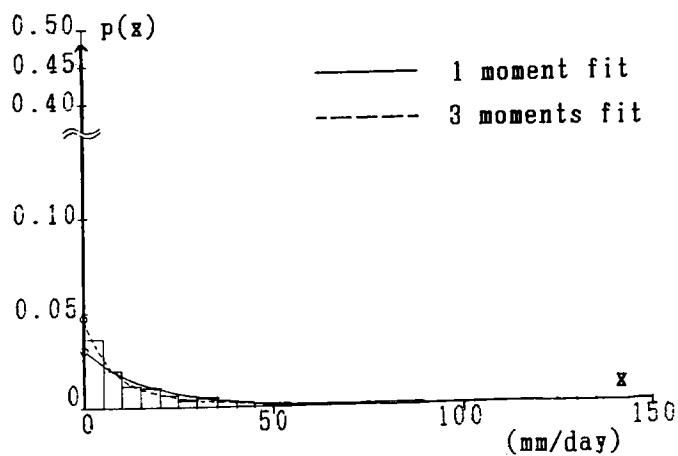


(a) Distribution with the pulse probability at the origin

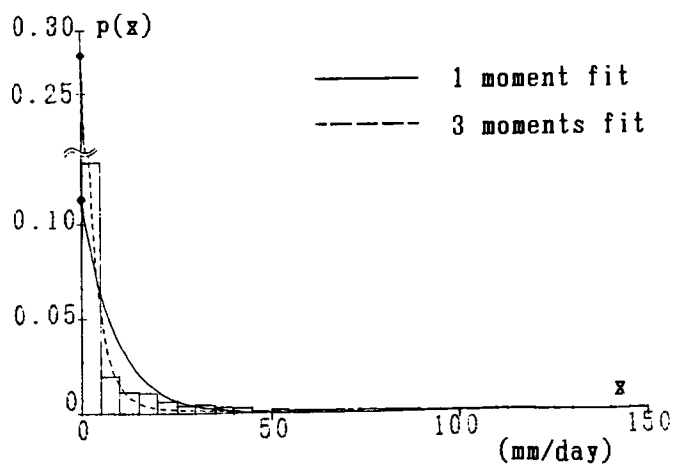


(b) Distribution with no pulse probability at the origin

図7.1 藪原日降水量のヒストグラムと最大エントロピー分布



(a) Distribution with the pulse probability at the origin



(b) Distribution with no pulse probability at the origin

図7.2 福島日降水量のヒストグラムと最大エントロピー分布

2.2 高次モーメント適合

我々が水文量の頻度解析をする上で知りたいのは母集団の確率密度関数であるが、それが未知であるため、限られたデータからいかにうまく母集団を推定するかということに多くの問題が集約される。それを大別すれば、確率密度関数の選択の問題とパラメタ推定の問題となる。後者の問題の1つに高次モーメントにおけるサンプリングエラーがある。しかし、そのモーメントも母集団に対する何らかの情報を有しており、最大エントロピー分布を利用する上で何次までのモーメントを用いるか、あるいはモーメントのもつ誤差をいかにおさえるかということ

表7.1 藪原日降水量から得られた最大エントロピー分布のパラメタの値

	λ_0	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
1 moment fit	1.06545	0.58700							
3 moments fit	0.73332	1.05764	-0.09686	0.00413					
5 moments fit	0.30432	3.09804	-1.87683	0.52801	-0.05974	0.00236			
6 moments fit	0.49947	1.79896	-0.20297	-0.25893	0.10310	-0.01275	0.00051		
7 moments fit	0.23162	4.15565	-4.38585	2.55461	-0.78534	0.12825	-0.01035	0.00032	
8 moments fit	0.84459	-2.69593	12.30450	-13.31000	6.58568	-1.70536	0.23766	-0.01678	0.00047

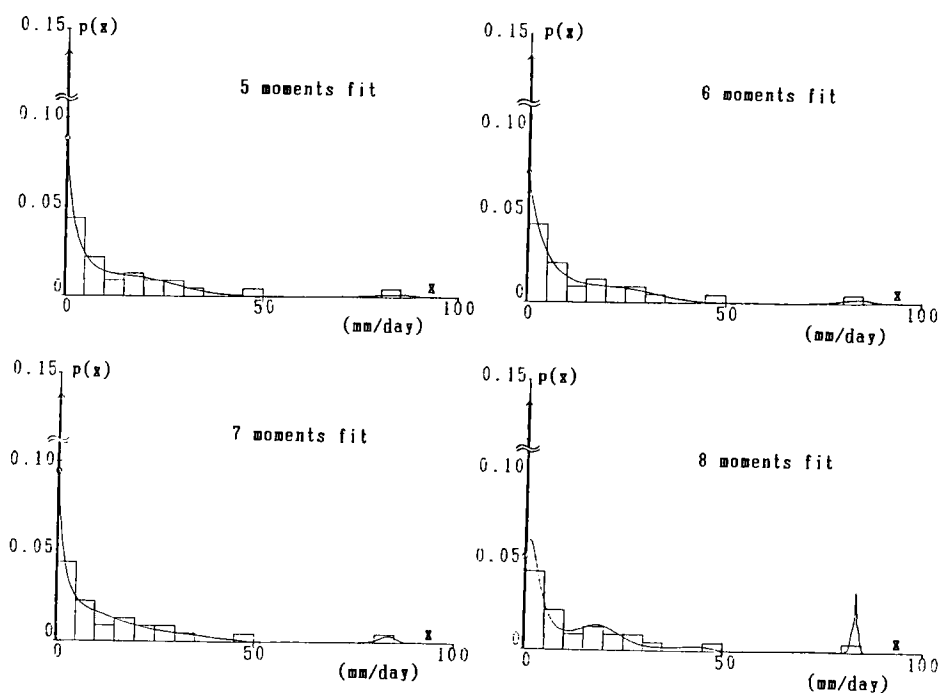


図7.3 藪原日降水量の高次モーメント適合

についての検討は重要と思われる。しかし、ここでは、この分布の究極の特性を調べるため、あえて高次モーメントまで採用して、ヒストグラムに対するこの分布の適合度を考察する。

表7.1 は、2.1.2 のパルスをもつ場合の藪原の日降水量について同定した推定分布のパラメタの値であり、図7.3 は、ヒストグラムと5～8モーメント適合との対応を示したものである。推定分布は、採用するモーメントの個数が増加するに伴い、ヒストグラムの細部にわたってよく記述できることがわかる。図7.3 の計算例では、7，8次モーメント適合になると、右端に離れて存在するヒストグラムの1部までよく追隨されている。

また、ここでは提示していないが、2.1.2 で述べた他の3つの計算例からも、同様な結果を得ている。

2.3 複数の極値をもつ分布への対応

最大エントロピー分布の特徴の1つは、複数の極値を表現できることにある。この性質はGram-Charlier の級数³⁾にもみられる。しかし、黒田・横田²⁾が示したようにGram-Charlier の級数では負の領域をとることがあり、確率密度関数の条件を満たさない場合が発生する。本項では、理論式と人工的に作成したヒストグラムに対する適合から、最大エントロピー分布が、確率密度関数の条件を保持しながら、複数の極値をもつ母集団を記述できることを示す。

2.3.1 理論的検討

一般的な多変数最大エントロピー分布は、式(4.5)で表現された。いま、これを x_i で偏微分して零とおくと

$$\frac{\partial p(x_1, x_2, \dots, x_n)}{\partial x_i} = \exp\left\{-1 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \cdot \left\{-\sum_{r=1}^N \lambda_r \frac{\partial g_r(x_1, x_2, \dots, x_n)}{\partial x_i}\right\} = 0, i=1, 2, \dots, N \quad (7.8)$$

となる。ここで、 $\exp\left\{-1 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \neq 0$ であるから、式(7.8)は

$$\sum_{r=1}^N \lambda_r \frac{\partial g_r(x_1, x_2, \dots, x_n)}{\partial x_i} = 0 \quad (7.9)$$

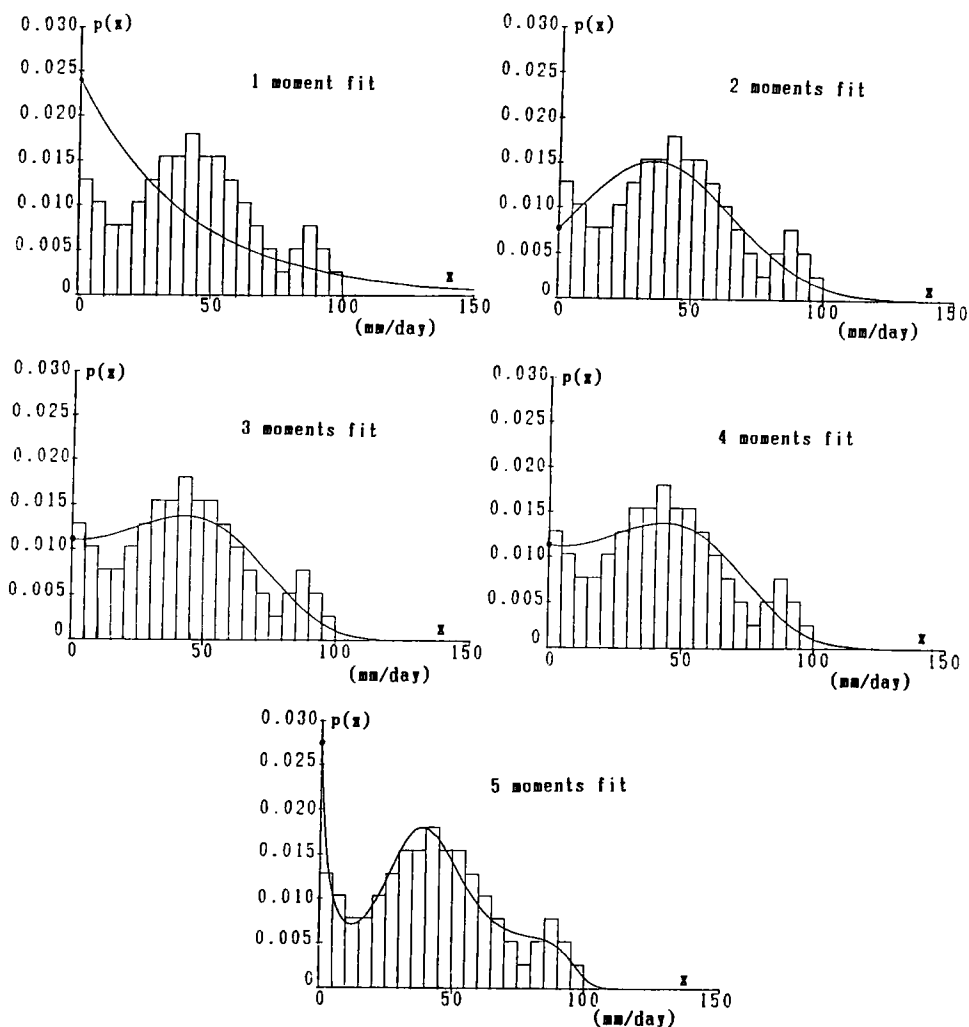


図7.4 人工的なヒストグラムと最大エントロピー分布（その1）

表7.2 人工的なヒストグラムから得られた
最大エントロピー分布のパラメタの値

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
1 moment fit	-0.000020	1.000010				
2 moments fit	1.142890	-1.646760	0.982830			
3 moments fit	0.768556	0.071836	-0.759612	0.473096		
4 moments fit	0.747348	0.224962	-1.016750	0.624380	-0.028642	
5 moments fit	-0.139149	11.656400	-32.819100	34.306900	-15.094500	2.39487

と等価である。上式は、 $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ を固定したときの x_i 方向の極値を与える。ここで、制約条件として $g_r(x)=x^r$ を採用するなら、文献2)にも述べられているように、 s 個の極値をもつ母集団を記述するためには $r \geq s+1$ とする必要がある。

また、式(4.5) は負の値をとらないため、上述の Gram-Charlierの級数で発生した負領域をとるという現象が、最大エントロピー分布では決して起こり得ないことは、明らかである。

2.3.2 ヒストグラムへの適用

降雨、流量の水文量ヒストグラムでは複数の極値が表われることはまれであるが、最大エントロピー分布の挙動を把握する上で、複数の極値をもつヒストグラムへの適合を考察することは重要と思われる。

ここでは、極値の状態を明確にするために、図7.4 のように人工的にヒストグラムを作成し、これに対してモーメントを制約条件とした1変数最大エントロピー分布の適用を試みた。表7.2 は、そのとき得られたパラメタの値を、図7.4 は1～5モーメント適合で得られた分布

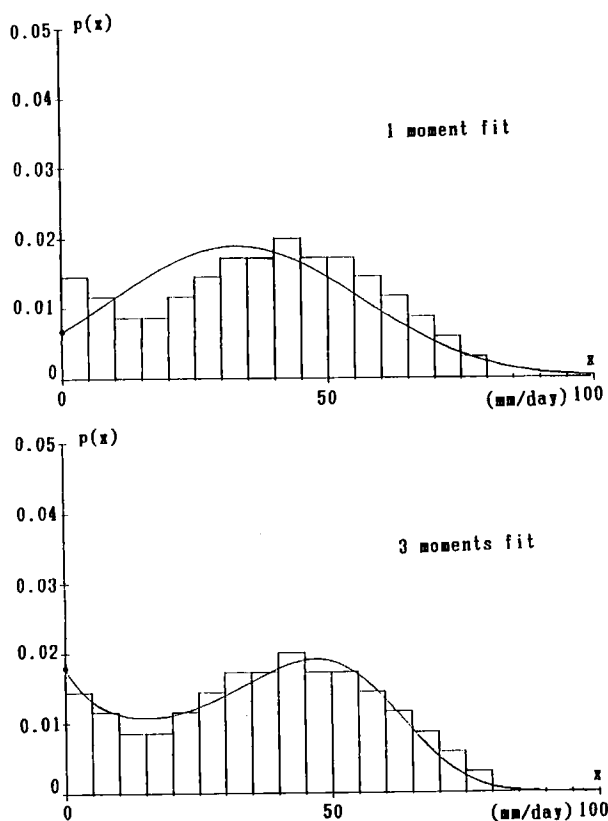


図7.5 人工的なヒストグラムと最大エントロピー分布（その2）

曲線を示している。ヒストグラムは4個の極値を持つため、5モーメント適合はよくヒストグラムを記述していることがわかる。

また、参考までに、図7.5に2つの極値をもつ人工的に作成したヒストグラムと2, 3モーメント適合の分布曲線を掲げておく。3モーメント適合が、ヒストグラムによく合致している。

第3節 要 約

本章では、前章までに議論しなかった、最大エントロピー分布が持つ特有の性質を3項目にわたって検討した。以下に、得られた成果を要約する。

1) 原点にパルス確率をもつ最大エントロピー分布を定式化し、日降水量のヒストグラムに適用して、第2章の最大エントロピー分布と比較した。その結果から、パルスを与えた場合、適合度の改善がみられ、その有効性が読み取れることを示した。

2) 水文統計では、一般に、4次までの統計モーメントしか用いられていないが、あえて最大8次までの統計モーメントを情報とした最大エントロピー分布を求め、日降水量のヒストグラムに適用した。適合結果から、推定分布は、採用するモーメントの個数が増加するに伴い、ヒストグラムの細部にわたってよく記述できるようになり、前節で取り上げた計算例では、7, 8モーメント適合になると、右端に離れて存在するヒストグラムの1部までよく追従できることを明らかにした。

3) 任意関数の期待値を制約条件とした、一般的な最大エントロピー分布の極値を表現する式を求めた。また、極値の状態が明確に表われる人工的なヒストグラムを作成し、それに統計モーメントで情報を与えた最大エントロピー分布を適用すると、理論どおりに極値を表現する推定分布が得られることを示した。

さらに、この後に示す Appendix Aは、いままで最大エントロピー法により推定してきた確率分布が、与えられた制約条件の下でエントロピーを最大にしていることを証明するものであり、Appendix Bは、与えられた制約条件の下でエントロピーを最大にするパラメタの組は唯一組であること、すなわち最大エントロピー分布のパラメタは一意に定まることを証明するものである。両者ともパラメタを同定する上で貴重な知識であるので、紹介しておく。

Appendix

黒田・横田²⁾は，原点のまわりの統計モーメントを制約条件とした離散的1変数最大エントロピー分布について，推定された確率密度関数がエントロピーを最大にしていること，および，その分布のパラメタの一意性を証明した。ここでは，任意関数を制約条件とした連続的多変数最大エントロピー分布についても同様なことが言えることを紹介し，第2章から第7章で同定したパラメタが上記の2つの条件を満足していたことを確認する。

Appendix A 推定された密度関数のエントロピー最大性

最大エントロピー法により推定された $p(x_1, x_2, \dots, x_n)$ と $g_r(x_1, x_2, \dots, x_n)$ の期待値が等しい確率密度関数を $p'(x_1, x_2, \dots, x_n)$ とする。すなわち

$$\begin{aligned} & \int \cdots \int g_r(x_1, x_2, \dots, x_n) p'(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \\ &= \int \cdots \int g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \end{aligned} \quad (1)$$

を仮定する。一般に，対数関数の上に向っての凸性から

$$\begin{aligned} & - \int \cdots \int p'(x_1, x_2, \dots, x_n) \ln p'(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \\ & \leq - \int \cdots \int p'(x_1, x_2, \dots, x_n) \ln p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \end{aligned} \quad (2)$$

が成立する。⁴⁾ 上式の右辺は

$$\int \cdots \int p'(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = 1 \quad (3)$$

と，式(4.5)，(1)を用いて次のように変形される。

(2)式の右辺

$$\begin{aligned}
 &= - \int \cdots \int p'(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \ln \left\{ \exp \left\{ -1 - \sum_{r=1}^N \lambda_r g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \right\} \right\} d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n \\
 &= 1 + \sum_{r=1}^N \lambda_r \int \cdots \int g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) p'(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n \\
 &= 1 + \sum_{r=1}^N \lambda_r \int \cdots \int g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) p(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n \\
 &= - \int \cdots \int p(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \ln p(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n \quad (4)
 \end{aligned}$$

ここで、 $p'(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ は式(1) を満足する任意の確率密度関数であるから、推定された $p(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ はエントロピーを最大にすることがわかる。

Appendix B パラメタの一意性

1 つの推定パラメタの組 $\lambda_r (r=1, 2, \cdots, N)$ とは別の、もう 1 つの推定パラメタの組を $\lambda'_r (r=1, 2, \cdots, N)$ とすると、このとき最大エントロピー分布は式(4.5) から

$$p(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) = \exp \left\{ -1 - \sum_{r=1}^N \lambda'_r g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \right\} \quad (1)$$

で与えられる。したがって、上式のエントロピー $H'(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ は

$$\begin{aligned}
 H'(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) &= - \int \cdots \int \left\{ -1 - \sum_{r=1}^N \lambda'_r g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \right\} \\
 &\quad \cdot \exp \left\{ -1 - \sum_{r=1}^N \lambda'_r g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \right\} d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n \\
 &= 1 + \sum_{r=1}^N \lambda'_r E[g_r(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)] \quad (2)
 \end{aligned}$$

となる。同様に $\lambda_r (r=1, 2, \dots, N)$ を推定パラメタの組とする最大エントロピー分布のエントロピー $-H(x_1, x_2, \dots, x_n)$ は

$$H(x_1, x_2, \dots, x_n) = 1 + \sum_{r=1}^N \lambda_r E[g_r(x_1, x_2, \dots, x_n)] \quad (3)$$

で与えられる。条件より $H(x_1, x_2, \dots, x_n)$ と $H'(x_1, x_2, \dots, x_n)$ は等しい。

すなわち

$$\sum_{r=1}^N \lambda_r' E[g_r(x_1, x_2, \dots, x_n)] = \sum_{r=1}^N \lambda_r E[g_r(x_1, x_2, \dots, x_n)] \quad (4)$$

であるから

$$\lambda_r' = \lambda_r, \quad r=1, 2, \dots, N \quad (5)$$

が成立する。故に、最大エントロピー分布のパラメタは一意に定まることがわかる。

参 考 文 献

- 1) 寒川典昭・荒木正夫・大原基秀・森川 智：最大エントロピー分布の諸特性の検討，第31回水理講演会論文集，pp.185～190，1987年 2月。
- 2) 黒田義輝・横田勝弘：最大エントロピー法による確率密度関数の推定法，計測自動制御学会論文集，第18巻，第 1号，pp.15～22，1982年 1月。
- 3) 例えば，佐藤良一郎：数理統計学（増補版），培風館，pp.186～201，1943年 5月。
- 4) 例えば，田中幸吉：情報工学，朝倉書店，pp.170～171，1969年 9月。

第8章 最大エントロピーヒストグラム

第1節 概 説

前章までの最大エントロピー分布は、密度関数の形状が滑らかな曲線を呈していた。ヒストグラムから判断して、水文量の母集団もそのようであることが予想されるから、今までの最大エントロピー分布の曲線が望ましいことは確かである。しかし、本章では、このような最大エントロピー分布が求まらないときに対処するために、各々のブロックの面積が等しいヒストグラムの形で提案された、エントロピーを最大にする分布の導入を検討する。

例えば式(2.36)、(2.37)で見られるように、統計モーメントで情報を与えた1変数最大エントロピー分布には、その存在範囲が示されている。2変数以上の場合、および他の関数で情報を与えた場合については、今のところ同様な存在範囲が明らかにされていないが、またそれを得るのは難解と思われるが、著者の数値計算の経験からすると、情報の状態によってはパラメタ同定がなされていない場合があった。このようなとき、情報の与え方を変更したり、情報量を減少させたりすれば、存在範囲内に最大エントロピー分布を持ち込むことができるようになる場合が多い。しかし、そうした対処の仕方は、水文統計の精度低下を招く恐れがあるので、でき得れば避けた方が望ましいと考えられる。

本章では、統計モーメントで情報を与えた場合について、この問題を解決するために、Collins-Wragg¹⁾によって定式化された最大エントロピーヒストグラム(別名c-ヒストグラム(Common-area histogramsの略))の水文統計への導入をはかる。²⁾ この分布は、母集団を有限個の、各々のブロックの面積が等しいヒストグラムで近似し、エントロピーを最大にするように個々のブロックの間隔を設定するものである。従って、ブロックの数を大きくすれば、滑らかな分布に近づくことが期待される。まず、第2節では、c-ヒストグラムの定式化とパラメタ同定法について紹介する。この分布を水文統計解析に導入するための思想的な背景は、本章2.1でも触れているように、第2章2.1で詳しく議論したことと同様である。第3節では、第2節の理論を、日降水量、月降水量、および年最大日降水量に適用して、c-ヒストグラムの特徴を最大エントロピー分布との比較において論ずる。特に、最大エントロピー分布が存在しない場合の水文データへのc-ヒストグラムの適用性、両者とも存在する場合の分布形状やリターン・ピリオドの比較および、パラメタ同定に要した計算時間の比較を通じて、両者の特徴を明らかにしていく。

なお、著者は今のところ、このヒストグラムは、前章までの最大エントロピー分布が存在し

ないときに、水文統計に活用する場面がでてくるのであり、基本的には今までの最大エントロピー分布を用いることを考えている。

第2節 c-ヒストグラムの推定

2.1 定式化

いま、確率変数 $x[0, R]$ 上に M 個のブロックを考え、 n 番目ブロックの確率密度を

$$w(x) = w_n = \frac{1}{Ma_n}, \quad x_{n-1} < x < x_n \quad (8.1)$$

で定義する。ただし

$$a_n = x_n - x_{n-1}, \quad n = 1, 2, \dots, N \quad (8.2)$$

とする。このとき、式(8.1)のエントロピーは、式(2.1)から

$$\begin{aligned} H_M &= - \int_0^R w(x) \ln w(x) dx \\ &= - \sum_{n=1}^M \int_{x_{n-1}}^{x_n} w_n \ln w_n dx \\ &= \ln M + \frac{1}{M} \sum_{n=1}^M \ln a_n \end{aligned} \quad (8.3)$$

となる。一方、式(2.28)に対応した原点のまわりの r 次のモーメントは

$$\begin{aligned}\mu_r &= \int_0^R x^r w(x) dx \\ &= \frac{1}{M(r+1)} \sum_{n=1}^M \sigma_{rn}\end{aligned}\quad (8.4)$$

で与えられる。ここに

$$\sigma_{rn} = x_{n-1}^r + x_{n-1}^{r-1} x_n + \cdots + x_{n-1} x_n^{r-1} + x_n^r, \quad x_0=0, \quad r=1,2,\dots,M \quad (8.5)$$

である。ただし、式(2.27)に対応した確率密度関数が具備すべき条件は、式(8.1)に含まれていることに注意を要する。

Collins・Wragg¹⁾ は、式(8.4)の制約の下で、式(8.3)を最大にする a_n から式(8.1)の w_n を求め、それをc-ヒストグラム(Common-area histogram)と呼んだ。

いま、式(8.4)の制約をはずして、簡単のために $M=2$ の場合を考えてみよう。このとき、式(8.3)は

$$H_2 = \ln 2 + \frac{1}{2} \ln(a_1 a_2) \quad (8.6)$$

となり、 $x[0,R]$ とするなら、上式を最大にする a_n ($n=1,2$) は

$$a_n = \frac{R}{2}, \quad n=1,2 \quad (8.7)$$

で与えられる。したがって、式(8.1)より

$$w_n = \frac{1}{R}, \quad n=1,2 \quad (8.8)$$

となり、これは $[0,R]$ の上での一様分布を表現している。

このことから類推して、本項でも第2章2.1の議論を展開することができる。すなわち、水文量の頻度分析に式(8.4)を制約条件として式(8.3)を最大にする、いわゆる Collins・Wragg

のc-ヒストグラムを導入することは妥当と判断される。

2.2 パラメタ同定法

本項でも、第2章2.4と同様な考え方を用いた、パラメタ同定過程が簡明な、Collins・Wragg¹⁾の手法を採用する。

前節で定式化された最大化問題は、第2章2.4で示したのと同じように、ラグランジュ関数

$$L = H_M + \sum_{r=1}^N \lambda_r \left\{ \mu_r - \frac{1}{M(r+1)} \sum_{n=1}^M \sigma_{rn} \right\} \quad (8.9)$$

の最大化を考えることにより解かれ、その解は

$$F_n = \frac{\partial H_M}{\partial x_n} = 0, \quad n=1, 2, \dots, M, \quad (8.10)$$

$$F_{M+r} = \mu_r - \frac{1}{M(r+1)} \sum_{n=1}^M \sigma_{rn} = 0, \quad r=1, 2, \dots, N \quad (8.11)$$

の非線形方程式を満足するように決定される。

いま、式(8.10)、(8.11)を線形化するため、解の近似値を

$$x_0 = \{x_1^0, x_2^0, \dots, x_M^0; \lambda_1^0, \lambda_2^0, \dots, \lambda_N^0\} \quad (8.12)$$

残差を

$$E = \{e_1, e_2, \dots, e_M; e_{M+1}, e_{M+2}, \dots, e_{M+N}\} \quad (8.13)$$

と置き、近似値のまわりで Taylor 展開し、 $e_i (i=1, 2, \dots, M+N)$ に関して1次以下の項のみ採用すると、次の方程式が得られる。

算結果から推察して零に設定している。

第3節 実データへの適用

本節では、前節の理論を実データに適用して、c-ヒストグラムの特徴を検討する。

実データとして、長野県内の観測所の次の降水量を採用した。

- 1) 福島：1938年～1983年までの9月15日の日降水量。
- 2) 蕨原：1938年～1983年までの6月15日の日降水量。
- 3) 長野：1888年～1985年までの6月の月降水量。
- 4) 長野：1890年～1973年までの年最大日降水量。
- 5) 松本：1898年～1969年までの年最大日降水量。

1)と2), 4)と5)は同じ時間単位の降雨であるため、図 8.1, 8.3, 8.5に 1), 3), 4) のヒストグラムを、図 8.2, 8.4, 8.6に、用いたモーメントとブロック数ごとのc-ヒストグラムおよび1変数最大エントロピー分布を示す。ただし、図中に最大エントロピー分布の掲載がない箇所は、パラメタ同定がなされなかった場合である。また、表 8.1に、参考までに図8.6(f)を求めたときの x_0 の初期値と解、および $w(x)$ の値を示しておく。

いま、得られた結果の特徴を、採用した降水の時間単位別に考察すると、次のことが言える。

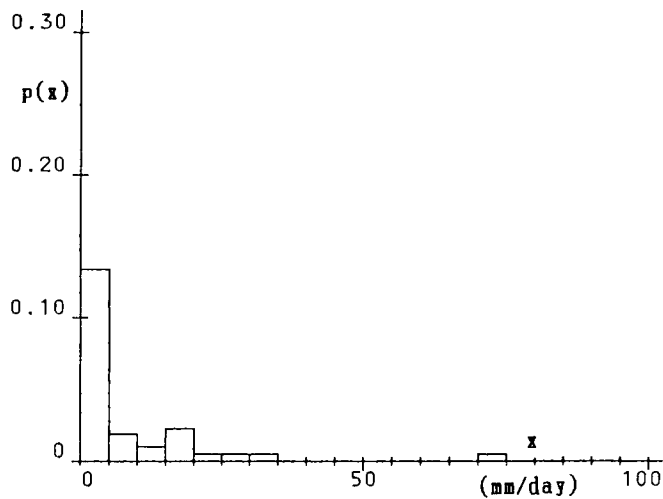
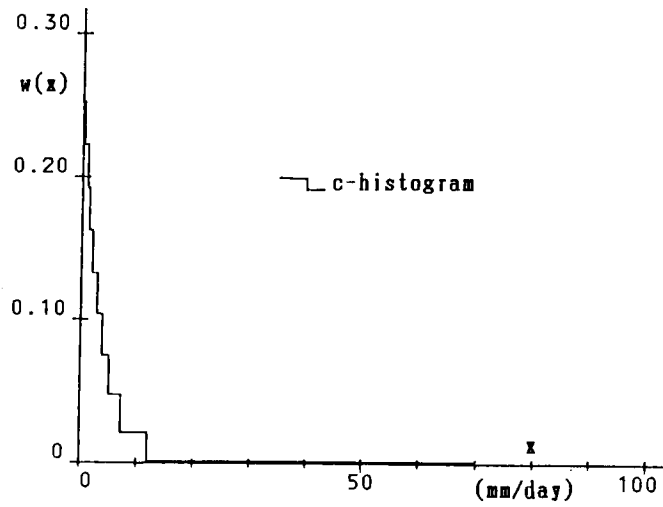
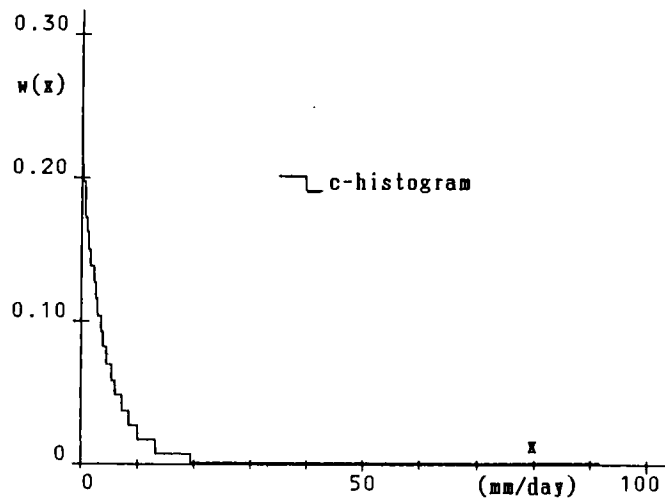


図8.1 福島 9月15日降水量のヒストグラム



(a) 2 moments fit, $M=10$



(b) 2 moments fit, $M=20$

図8.2 福島 9月15日降水量のc-ヒストグラム

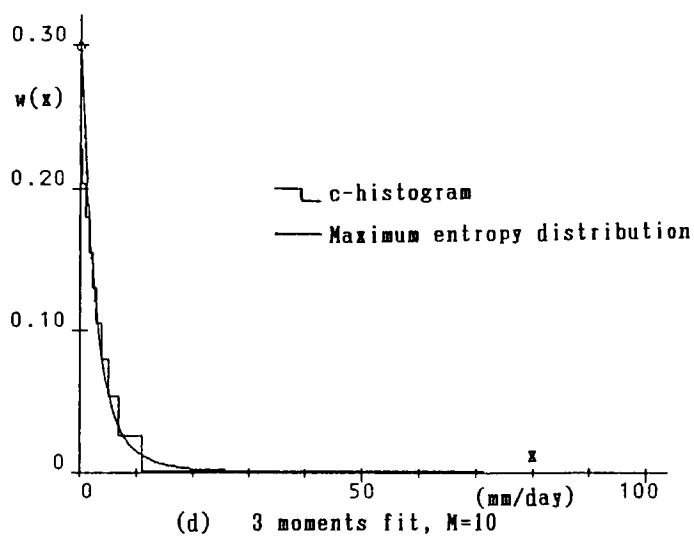
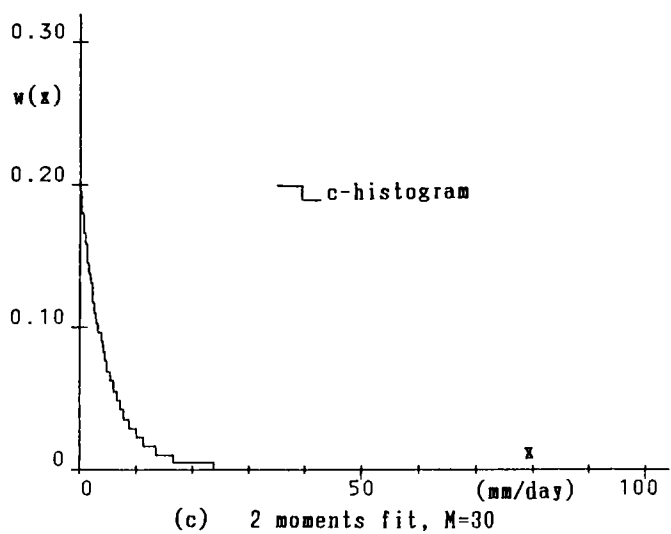


図8.2 (続き)

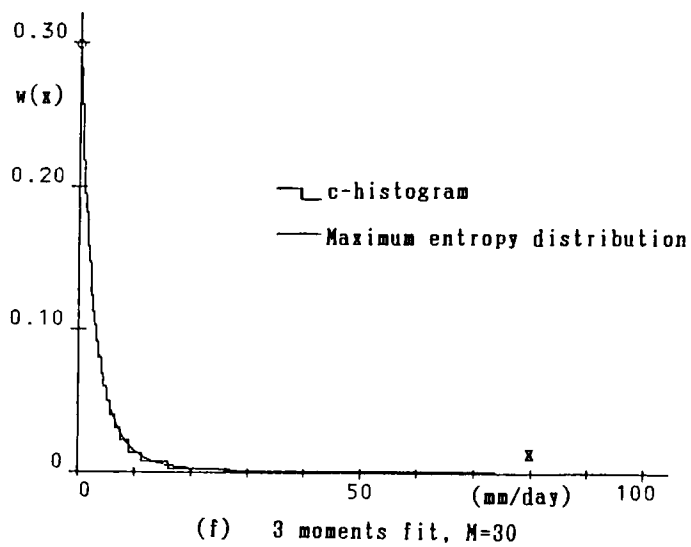
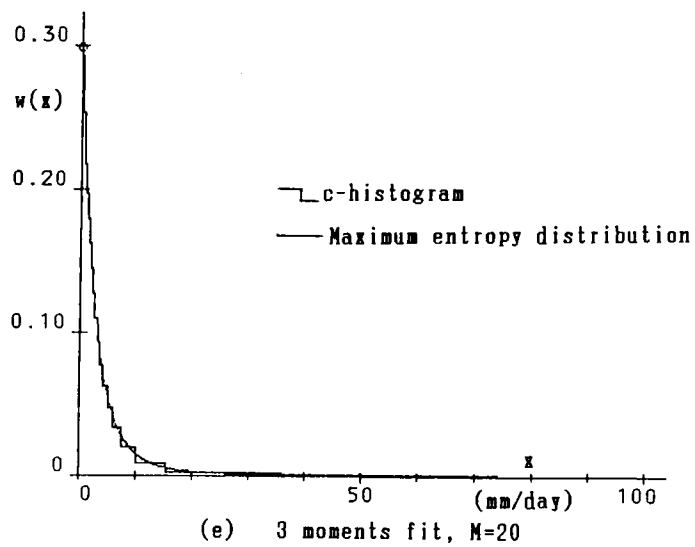


図8.2 (続き)

a) 1), 2) の降水量

μ_1 , μ_2 を原点のまわりの 1 次, 2 次モーメントとすると, 1) の場合, $\mu_1=6.9867$, $\mu_2=211.36$, 2) の場合, $\mu_1=11.3889$, $\mu_2=891.314$ であるから, どちらも $2\mu_1^2 < \mu_2$ となり, 式(2.39)の存在条件を満足していないことになる。しかし, この場合も, 2 モーメント適合の c-ヒストグラムは, 図 8.2(a) ~ (c) のように求まることがわかる。ここで, 採用した日降水量の場合は, $M=30$ の 3 モーメント適合で, 全体的にも, 原点付近および右裾の方も十分にヒストグラムに適合しており, M の値を 30 にとれば, 最大エントロピー分布に対する近似度も十分であった。

b) 3) の降水量

c-ヒストグラムによる 2 モーメント適合の実データヒストグラムに対する適合度は, 前者の方がフラットになりあまりよくない。しかし, 採用するモーメントの個数を増加させるに従ってしだいに改善され, $M=30$ で 4 次モーメントまで採用すると, 全体的に十分な適合度が得られるとともに, 240mm 程度を越える点での確率密度の値が大きくなり, 左裾, 右裾に対してもよく適合している。また, c-ヒストグラムと最大エントロピー分布の傾向は, ほぼ一致しているが, M の値は, 最大エントロピー分布への近似度合から判断しても, 30 程度とする必要があった。ただし, 4 モーメント適合の最大エントロピー分布は求まらず, 同じく c-ヒストグ

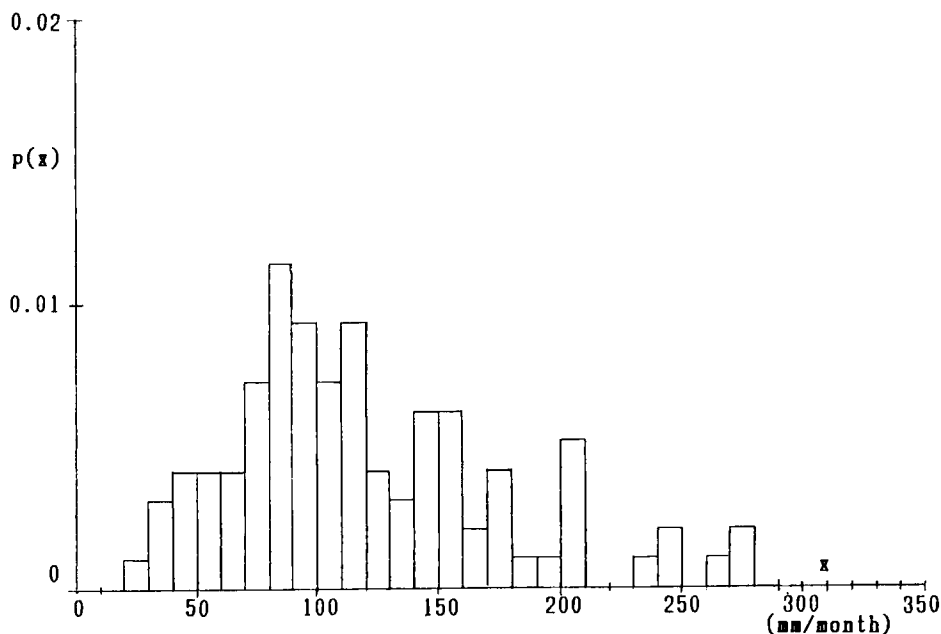


図8.3 長野 6月降水量のヒストグラム

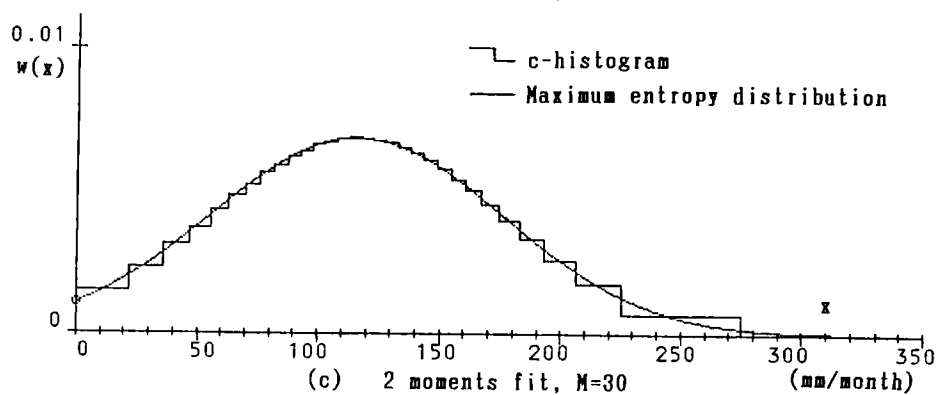
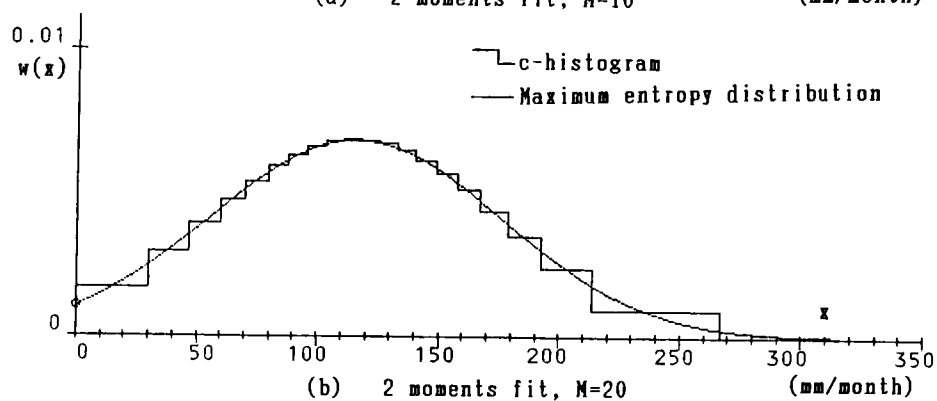
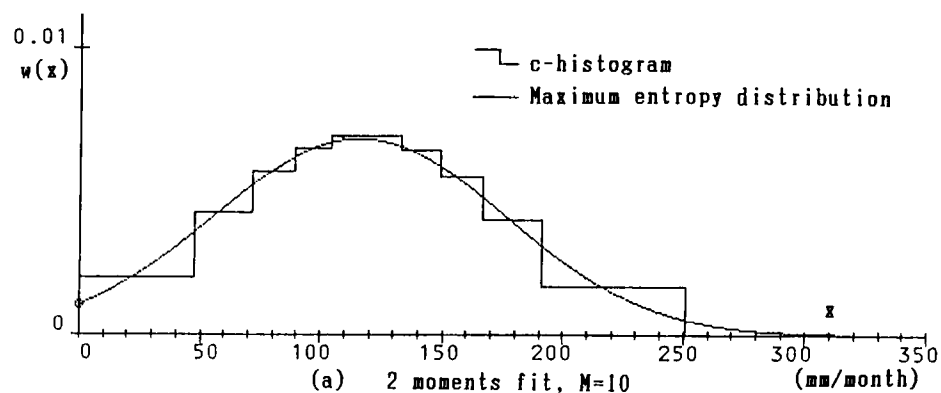


図8.4 長野 6月降水量のc-ヒストグラム

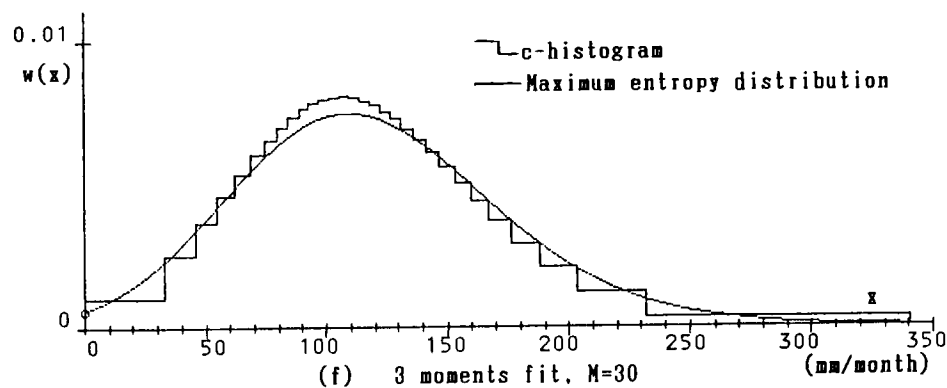
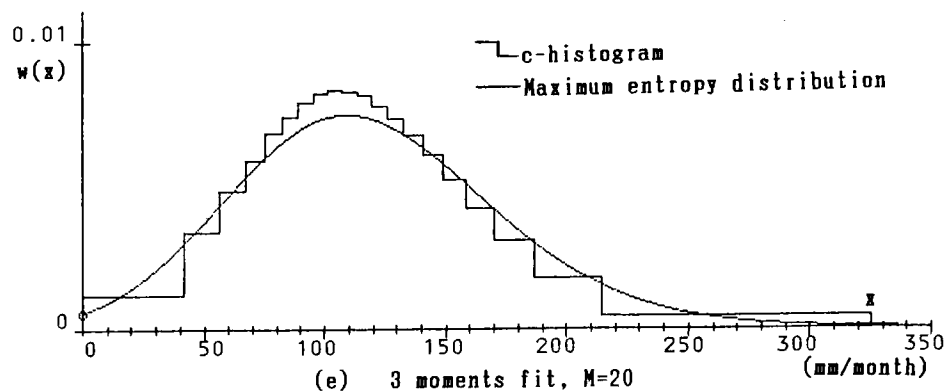
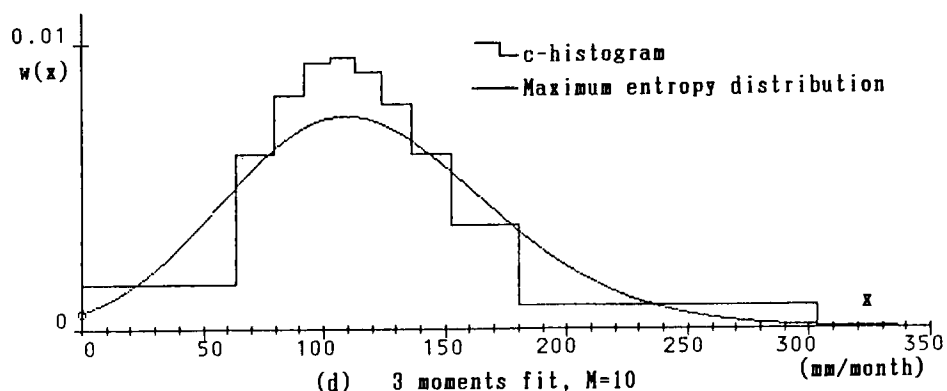


図8.4 (続き)

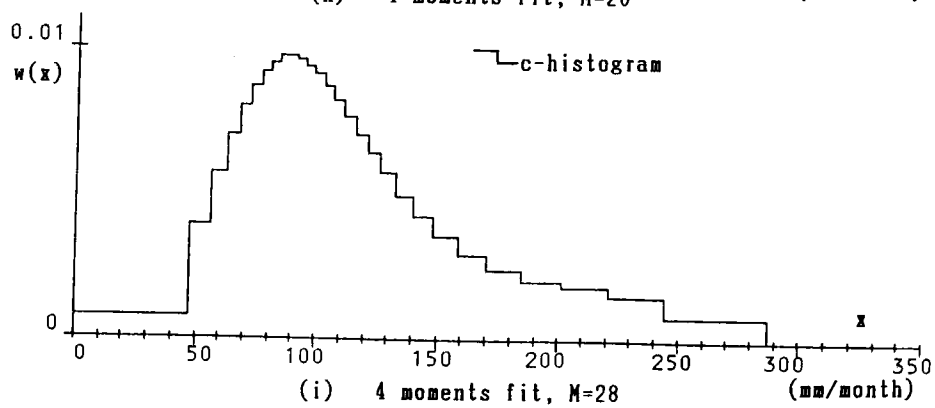
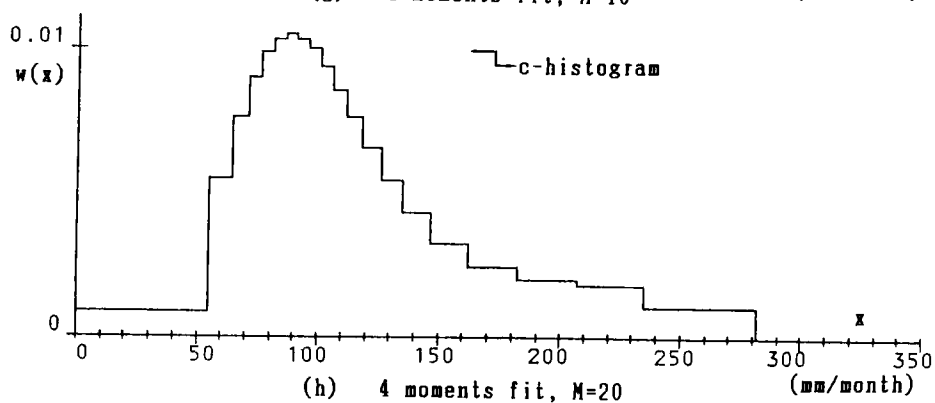
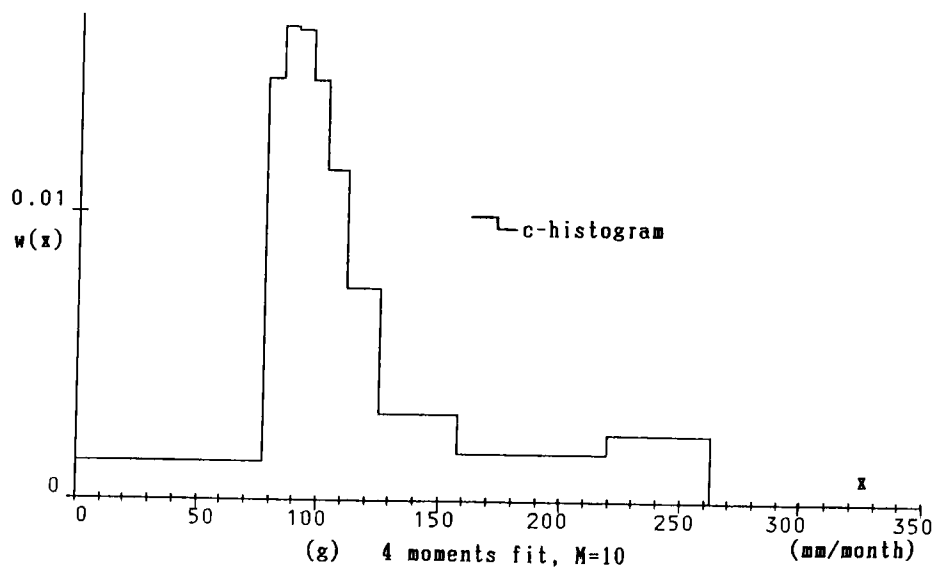


図8.4 (続き)

ラムは $M=29$, 30の場合は得られなかったため, $M=28$ としている。

c) 4), 5) の降水量

この場合も, 2モーメント適合のc-ヒストグラムは実データのヒストグラムよりフラットになって, あまり適合度はよくないが, 3モーメント適合にすると, 分布形にシャープ度が変わり, ピークの位置も左にずれて, 実データのヒストグラムに近づくとともに, 右裾に長く尾を引くようになり, ここで取り上げたような極値水文量を対象とした場合, 図8.6(f)に示した程度により大きな水文量の出現の可能性を与えることは好ましい傾向である。また, M の値は, a), b) の場合と同様, 30程度必要と考えられる。一方, 最大エントロピー分布は, 3次モーメントを加えてもほとんど分布形に変化がなく, 3モーメント適合ではc-ヒストグラムの方が最大エントロピー分布よりシャープになって, 実データヒストグラムへの適合もよくなっている。また, 表 8.2は, 右裾部分の実用的な適合度をみるために, データ中の最大値のリターン・ピリオドについて, 最大エントロピー分布とc-ヒストグラムの間で比較したものである。3モーメント適合の場合, c-ヒストグラムの方が最大エントロピー分布よりかなり小さな値を示している。これは, 図8.6(d)~(f) からわかるように, 3モーメント適合にすると3次モーメントの導入により, c-ヒストグラムが最大エントロピー分布より右裾に長く尾を引くためである。なお, ここで取り上げた降水量では, 4モーメント適合は, 式(8.12)で示し

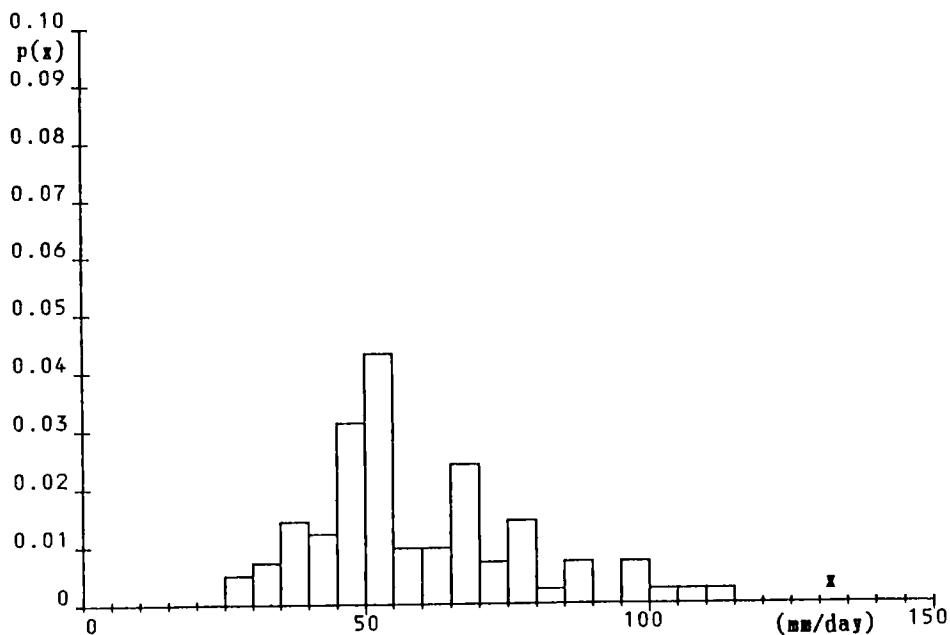


図8.5 長野年最大日降水量のヒストグラム

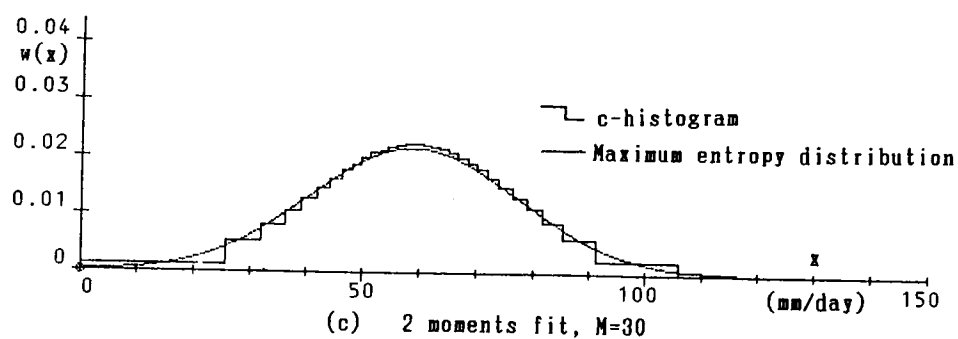
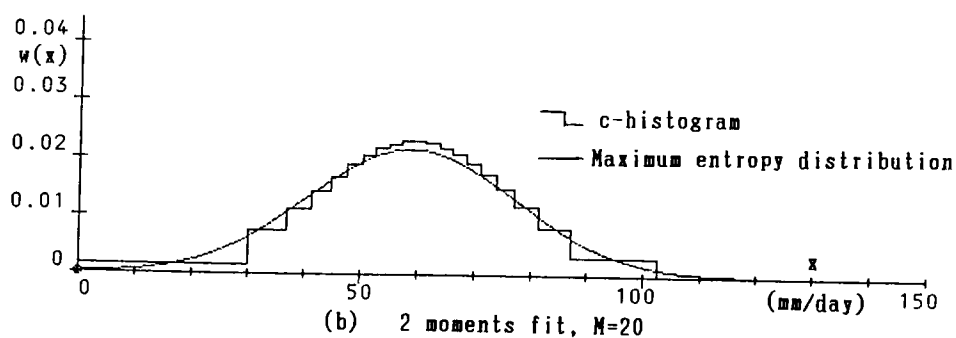
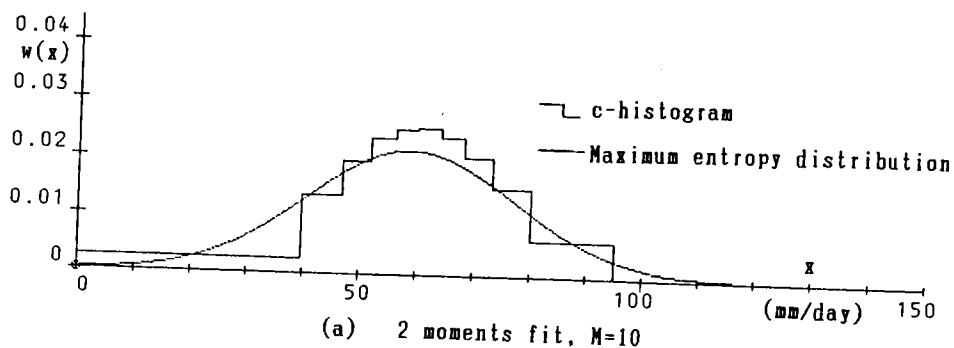


図8.6 長野年最大日降水量のc-ヒストグラム

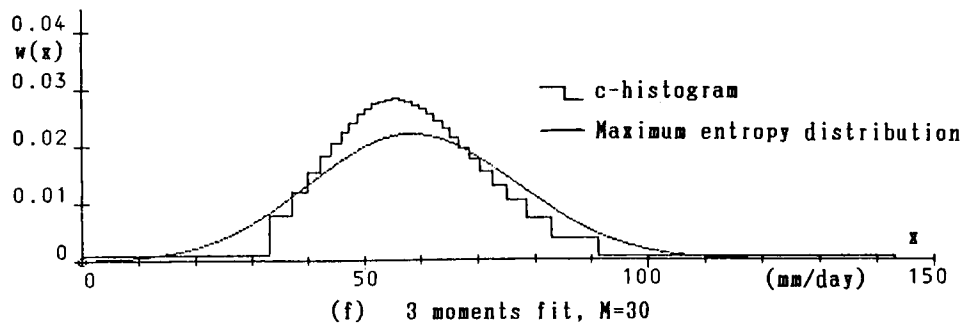
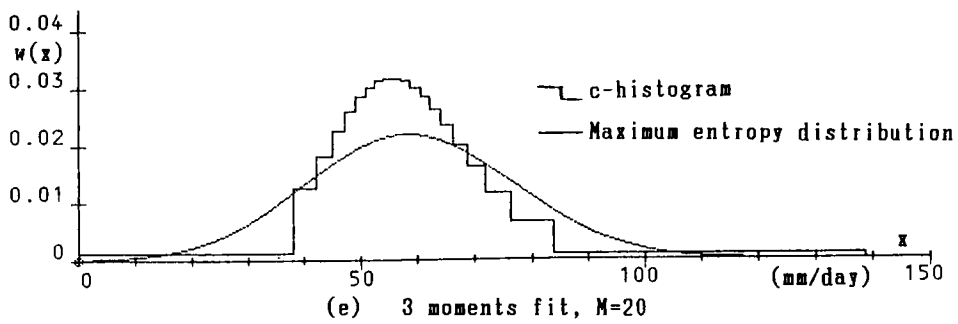
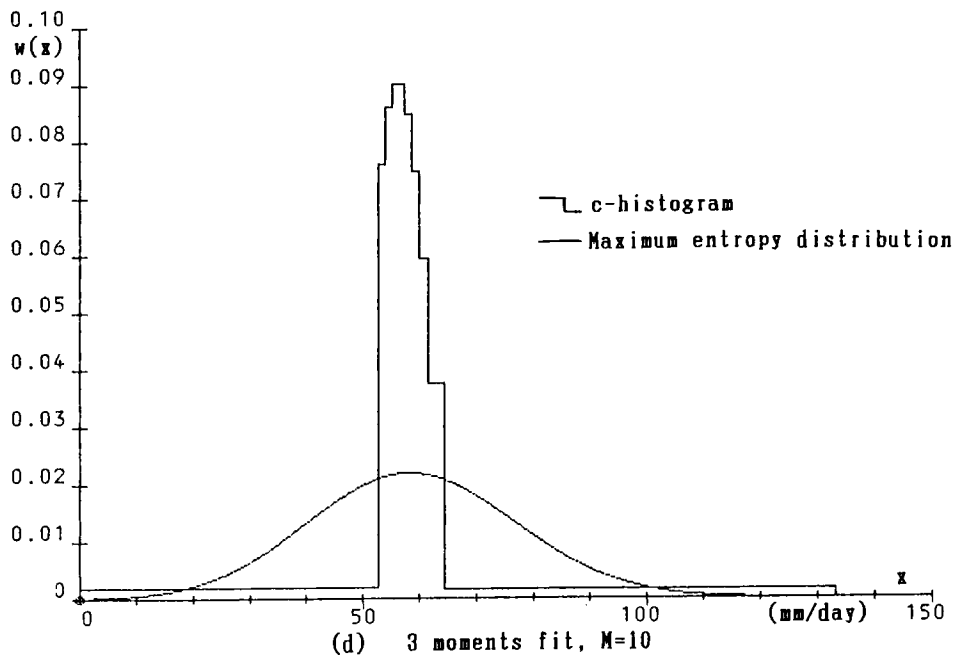


図8.6 (続き)

表8.1 図8.6 (f) を求めたときの
初期値, x_n , および $w(x)$ の値

Initial value	x_n	$w(x)$
30.0	32.63	0.001022
40.0	45.63	0.002515
50.0	55.04	0.003642
60.0	62.34	0.004564
70.0	68.59	0.005335
80.0	74.17	0.005981
90.0	79.28	0.006521
100.0	84.06	0.006968
110.0	88.61	0.007329
120.0	92.99	0.007612
130.0	97.25	0.007822
140.0	101.44	0.007962
150.0	105.58	0.008038
160.0	109.73	0.008050
170.0	113.89	0.008001
180.0	118.11	0.007893
190.0	122.43	0.007727
200.0	126.87	0.007504
210.0	131.48	0.007225
220.0	136.32	0.006889
230.0	141.45	0.006497
240.0	146.96	0.006049
250.0	152.98	0.005542
260.0	159.68	0.004977
270.0	167.34	0.004351
280.0	176.44	0.003663
290.0	187.90	0.002908
300.0	203.89	0.002085
350.0	231.88	0.001191
400.0	340.73	0.000306

た初期値 x_1^0 の設定がうまくいかず, 求まらなかった。

次に, 1)~5)の降水量に対して, パラメタ同定に要した計算時間を表 8.3に示した。3)~5)の3モーメント適合では, c-ヒストグラムの計算時間は, 最大エントロピー分布と比較して 1/5~ 1/9程度となっており, また, どの場合もブロック数の増加にあまり依存しないことがわかった。

以上の考察より, c-ヒストグラムを用いることの利点がいくつか明らかにされたが, この分布は, x_1 の初期値の設定に, 特に M の値が大きいきとき多くの試行錯誤を必要とすること, および2変数以上への拡張が今のところ難しいこと, 等の欠点を有していることを付記しておく。

表8.2 最大値のリターン・ピリオド (年)

(a) Annual maximum daily rainfall

(Nagano, Number of data:84)

Number of block		10	20	30
2 moments fit	C-H	∞	∞	∞
	MED	729.6		
3 moments fit	C-H	36	44	53
	MED	640.2		

(b) Annual maximum daily rainfall

(Matumoto, Number of data:72)

Number of block		10	20	30
2 moments fit	C-H	∞	∞	156
	MED	729.6		
3 moments fit	C-H	23	31	39
	MED	119.2		

C-H:c-histogram

MED:Maximum entropy distribution

表8.3 パラメタ同定に要した計算時間 (秒)

(信州大学大型計算機センター HITAC M-240H)

Rain guage station	Number of moment	MED	C-H		
			10 *	20 *	30 *
Fukushima	2		5.64	5.60	5.93
	3	8.94	7.19	7.84	8.95
Yabuhara	2		5.70	5.82	6.01
	3	8.74		7.99	9.15
Nagano	2	6.56	5.33	5.46	5.55
	3	38.88	7.22	7.25	8.08
Nagano Annual maximum daily rainfall	2	8.39	5.30	5.27	5.55
	3	67.86	7.34	7.52	7.94
Matumoto Annual maximum daily rainfall	2	8.45	5.52	5.69	5.81
	3	66.33	7.29	7.66	8.26

*:Number of block

第4節 要 約

本章では、滑らかな曲線を有する最大エントロピー分布が存在しない場合に対処するために、 c -ヒストグラムの水文統計への導入について考察した。

第2節は、主として、Collins-Wragg の提案した c -ヒストグラムの定式化とパラメタ同定法の紹介であるので、ここでは、第3節で得られた成果のみを以下に要約する。

1) 日降水量データに対する適用では、最大エントロピー分布の2モーメント適合の存在条件を満足していないデータに対して、 c -ヒストグラムが求まることを示した。また、このようなヒストグラムに対して、ブロック数を30程度にとるのが適当であることを述べた。

2) 月降水量データに対する適用では、4次モーメントまでを用いた c -ヒストグラムで、十分にデータのヒストグラムを記述できた。また、 c -ヒストグラムと最大エントロピー分布の傾向はほぼ一致していたが、実データのヒストグラムに対する適合、および最大エントロピー分布の近似度合から判断して、この場合もブロック数は30程度必要であることを述べた。

3) 年最大日降水量に対する適用では、3モーメント適合の場合、 c -ヒストグラムの方が、最大エントロピー分布よりシャープになって、実データヒストグラムへの適合もよくなっていることを述べた。また、データ中の最大値のリターン・ピリオドの値を両者の分布で比較すると、3モーメント適合の場合、 c -ヒストグラムの方が最大エントロピー分布よりかなり小さな値となることを述べた。

4) パラメタの同定に要した計算時間を、ここで用いた全降水データについて、 c -ヒストグラムと最大エントロピー分布の間で比較して、月降水量と年最大日降水量の3モーメント適合では、 c -ヒストグラムの方が $1/5 \sim 1/9$ 程度になっており、また c -ヒストグラムのパラメタ同定時間は、どの降水量の場合もブロック数にあまり依存しないことを明らかにした。

以上、いくつかの計算例に基いて得た c -ヒストグラムの特徴を要約したが、第1節でも記述したように、著者は、最大エントロピー分布の弱点を補う立場からこのヒストグラムの利用を考えている。

参 考 文 献

- 1) Collins, R. and Wragg, A. : Maximum entropy histograms, J. Phys. A : Math. Gen., Vol.10, No.9, pp.1441 ~1464, 1977.
- 2) 荒木正夫・寒川典昭・森川 智・柴山敏英：最大エントロピーヒストグラムの水文頻度分析への導入，土木学会中部支部研究発表会講演概要集，II-3, pp.106~107, 1987年3月。

第 9 章 結 論

水工計画における、確率分布による水文統計の要請は、1 変数の場合は言うに及ばず、2 変数、3 変数以上の多変数、および条件付きの場合においても非常に大きい。それに対処するための理論面での研究は、まだまだ不十分であった。本研究では、このような要請と、それらの分布の開発状況を踏まえて、上述のどの場合の分布にもなり得る最大エントロピー分布の開発、その特徴の分析、および実データへの適用を検討した。以下に、得られた研究成果を要約して結論としよう。

第 1 章では、水工計画を作成する上で、上述のそれぞれの変数の分布の重要性と、本研究以前の各々の場合の確率分布の状況について概説しながら、水文統計における最大エントロピー分布の必要性を述べ、本研究の目的とその内容を明らかにした。

第 2 章から第 5 章までは、変数の数の増加、変数の指定による最大エントロピー分布の拡張を取り扱った。

第 2 章では、1 変数最大エントロピー分布の理論を要約して提示した。特に、最大エントロピー分布による水文量の頻度分析の妥当性をこの分布の定式化を通じて明らかにするとともに、最大エントロピー分布が 1 つの分布系として位置付けられるべきであることを、統計学でよく知られている既存分布がこの分布から導出されていることを用いながら、Pearson 系の分布との比較において論じた。ついで、既存分布を母集団と仮定した場合と、実データを用いた場合について、統計モーメントで情報を与えた 1 変数最大エントロピー分布の適合度を検討して、一般には 4 次モーメントまでを情報とすれば十分であることを示した。また、この時用いた実データは、種々の時間単位の、多くの降雨・流量データであることを提示した。さらに、1 変数最大エントロピー分布の、モーメントの変動に伴う感度分析の結果も示した。

第 3 章では、2 変数最大エントロピー分布を提案した。定式化は一般的な情報の与え方で行った。また、ここでも、前章に準じて、2 変数 Gamma 分布を母集団と仮定した場合と、関連のある 2 地点で観測された実データを用いた場合について、統計モーメントで情報を与えた 2 変数最大エントロピー分布の適合度を検討した。得られた結果から、母集団が指数分布をする場合は 3 次モーメントまで、その他の場合は 4 次モーメントまで情報とすれば十分であることを示した。さらに、前章と同様に、モーメントの変動に伴う、2 変数最大エントロピー分布の感度を分析してその結果を提示した。一方、理論面では、2 変数最大エントロピー分布から、2 変数一様、正規、関連のない指数分布を導出できることを示したが、特に 2 変数正規分布と 2 変数指数分布の導出については、Appendix で詳述した。

第4章では、前章の2変数最大エントロピー分布を、3変数以上の多変数へと拡張した。この際、最大エントロピー分布の水文統計への導入の思想的な背景となった多変数の一様分布を導出しておいた。次に、3変数正規分布を母集団と仮定した場合と、相関のある3地点で観測された実データを用いた場合について、統計モーメントで情報を与えた3変数最大エントロピー分布の適合度を調べた。その結果から、前者の場合は、母集団が正規分布であるので、すべての2次以下のモーメントを情報とすると十分よい再現性が得られること、後者の場合は、全体的なヒストグラムの記述には2次モーメントまででよいが、さらに精度よくヒストグラムを説明するには4次モーメントまでを情報とする必要があることを示した。この分布により、非正規水文量の原型のままでの、確率分布による多変数統計が可能となった。

第5章では、前章の多変数最大エントロピー分布から、条件付き最大エントロピー分布の定義に従って、任意個数の変数を条件とし、他の任意個数の変数を対象とする条件付き最大エントロピー分布を提案した。さらに、条件の数、および制約するモーメントの個数を指定して、5個の条件付き最大エントロピー分布の理論式を導いた。また、統計モーメントで情報を与えた条件付き最大エントロピー分布の実データへの適用では、まず、年降水量を対象として、モーメントの取り方と条件の取り方による分布形の変化を調べた。この時、特に、相関構造の取り入れ方が分布の形状に大きく影響を及ぼすことを述べた。ついで、年最大日降水量を用いて、対象とした地点の降水量が、条件とした地点の降水量と同時に発生すると仮定する方法と、模擬降水を利用する方法の両者から、対象地点の降水量を推定した。得られた結果から、対象地点の推定値の、Weibull plotの方が、実測値の場合より小さな値を示していたことを述べた。

第6章と第7章では、最大エントロピー分布の特徴分析を取り扱った。

第6章では、情報の与え方と最大エントロピー分布の関係を論じた。ここでは、まず、最大エントロピー分布が、情報の与え方により統計学でよく知られている分布形の多くを記述できることは、対象とする水文量ごとに、適切な情報の与え方をみつけることが可能であるということを示唆していると考えた。そこで、まず情報を与える関数を $g_r(\cdot)$ として、4つの関数形を取り上げ、1変数Gamma分布を母集団と仮定した場合の適合度を検討した。このとき、母集団が指数分布の場合は、べき関数で情報を与えるのが良く、その他の場合は、べき関数と指数関数を組み合わせた情報が良いことを示した。さらに、べき関数を $g_m(x)=x^m$ 、指数関数を $g_n(x)=\exp(-nx/M)$ （ここで、 m, n ：正の整数、 $M=E[x]$ ）と表現すると、指数分布以外の母集団に対して、 $(m,n)=(1,2),(1,3),(1,4),(2,3),(2,4)$ の最大エントロピー分布がよく母集団を再現できることを明らかにし、特に、 $(m,n)=(1,3),(1,4)$ は変動が小さく、小さな値のリターン・ペリオドを持つことを示した。ついで、5地点で得られた年最大日降水量データを用いて、上述の5つの情報を持つ最大エントロピー分布で母集団推定を行って、どの分布もデータ

のヒストグラムに対してよい適合度を示すとともに、リターン・ピリオドも比較的安定して妥当な値を与えることを明らかにした。

第7章では、最大エントロピー分布の多くの特徴を把握しておくことを目的に、前章まで議論されなかった諸特性を検討した。まず第1は、原点にパルスを持つ最大エントロピー分布の提案と、日降水量ヒストグラムへの適用であり、パルスを持たない最大エントロピー分布と比較して、適合度が改善されることを示した。ついで、前章までは、情報として4次モーメントまでしか用いなかったが、ここでは、あえて、最大8次モーメントまでを情報とした最大エントロピー分布を求め、日降水量へ適用を試みた。得られた結果から、ここで取り上げた計算例では、7, 8モーメント適合の最大エントロピー分布になると、右端に離れて存在するヒストグラムの一部まで、よく追従できることを示した。最後に、複数の極値を持つヒストグラムへの最大エントロピー分布の対応について検討し、統計モーメントで情報を与えた最大エントロピー分布を用いると、理論どおり極値を表現する推定分布が得られることを明らかにした。また、パラメタ同定の基礎知識として必要であるため、Appendix Aでは、いままで最大エントロピー法で推定してきた確率分布が、与えられた制約条件の下でエントロピーを最大にしていること、Appendix Bでは、与えられた制約条件の下でエントロピーを最大にする確率分布のパラメタの組は唯一組であることを紹介した。

第8章では、エントロピーを最大にするヒストグラムを取り扱った。

すなわち、ここでは、滑らかな曲線を有する最大エントロピー分布が存在しない場合に、第2章の最大エントロピー分布の代用として、最大エントロピーヒストグラムの水文統計への導入を考察した。まず、理論式を要約して紹介した。ついで、この分布を日降水量、月降水量、年最大日降水量に適用して、最大エントロピー分布との比較において、この分布の特徴を述べた。特に、最大エントロピーヒストグラムは、最大エントロピー分布が得られない場合にも求まり、その代用となり得ること、最大エントロピーヒストグラムのパラメタ同定に要する計算時間は最大エントロピー分布と比べて短縮できる場合が多く、それはブロック数の増加にあまり依存しないこと、ブロック数は30個程度必要であること、を明らかにした。また著者は、いまのところ、最大エントロピー分布の弱点を補う立場からこの分布の利用を考えているという見解を示した。

以上述べたように、本論文は、確率分布を必要とする水文量の頻度分析のあらゆる場面に用いられる最大エントロピー分布の提案と、この分布の水文量への適用をめざした特徴分析という立場にたって、理論的、基礎的研究を行ったものである。本研究で得られた成果は、社会的要請に基づく合理的な水工計画を行う上で、大きく寄与するものと確信して結びとしたい。

謝 辞

本研究を遂行するにあたって、終始変らぬ御指導と御鞭達を賜りました京都大学工学部高棹琢馬教授，ならびに，京都大学防災研究所池淵周一教授に深く感謝し，心から御礼申し上げます。

また，本研究を進めるにあたり，数多くの御助言と激励をいただきました信州大学工学部荒木正夫教授，ならびに，富所五郎助教授に厚く御礼申し上げます。

さらに，京都大学工学部椎葉充晴助教授，岐阜大学工学部小尻利治助教授には，常に適切な水文学に関する御教示をいただきました。ここに深く感謝いたします。

最後に，信州大学工学部土木工学科の諸先生方，職員の方々には，日頃から温かい御支援を賜りました。同水工学講座に在籍された諸兄には，熱心な討論を通じて本研究に御協力頂きました。ここに，深く御礼申し上げ，感謝の意を表します。